
WQSRH
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U. K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
on: Wed Aug 16 09:55:04 2000; MasPar time 5.39 seconds
Tabular output not generated. 580.168 Million cell updates/sec

Title: >US-09-427-873-2
Description: (1-101) from US09427873.pep
Perfect Score: 101
Sequence: 1 LKFSQTCYNIAIQGSVLTS.....STKLNDDHIANIDGTLKYE 101

Scoring table: TABLE uniprotatable
Gap 60

Searched: 85661 seqs, 30989116 residues

Post-processing: Minimum Match 0%
Listing first 1000 summaries

Database: swiss-prot38
1:swissprot

Statistics: Mean 3.234; Variance 0.372; scale 8.703

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Alt	Score	Query Match	Length	ID	Description	Pred. No.
1	60	59.4	101	1 CVN_NOSEL	CYANOVIRIN-N (CV-N)	5.03e-177
2	8	7.9	678	1 DORS_DROME	EMBRYONIC POLARITY DOR	4.79e-04
3	7	6.9	259	1 PRG6_LYCES	PROTEASOME ALPHA SUBUN	8.39e-02
4	7	6.9	261	1 MODA_MYCTU	MOLYBDATE-BINDING PROT	8.39e-02
5	7	6.9	291	1 BHC1_RHOGO	BIPHENYL-2,3-DIOL 1,2-	8.39e-02
6	7	6.9	350	1 OSTG_YEAST	DOLICHYL-DIPHOSPHOOLIG	8.39e-02
7	7	6.9	440	1 PUR8_HELP1	ADENYLOSUCCINATE LYASE	8.39e-02
8	7	6.9	656	1 DNAK_ALCEU	DNAK PROTEIN (HEAT SHO	8.39e-02
9	7	6.9	879	1 YDBH_ECOLI	HYPOTHETICAL 96.8 KDA	8.39e-02
10	7	6.9	1113	1 PER3_MOUSE	PERIOD CIRCADIAN PROTE	8.39e-02
11	7	6.9	1148	1 ICEK_PSESY	ICE NUCLEATION PROTEIN	8.39e-02
12	7	6.9	1196	1 ICEV_PSESY	ICE NUCLEATION PROTEIN	8.39e-02
13	7	6.9	1200	1 ICEN_PSESY	ICE NUCLEATION PROTEIN	8.39e-02
14	7	6.9	1416	1 BLM_MOUSE	BLOOM'S SYNDROME PROTE	8.39e-02
15	6	5.9	84	1 Y076_NPVAC	HYPOTHETICAL 9.4 KDA P	8.75e+00
16	6	5.9	105	1 RL44_YEAST	60S RIBOSOMAL PROTEIN	8.75e+00
17	6	5.9	175	1 HER1_CAEEI	HER-1 PROTEIN PRECURSO	8.75e+00
18	6	5.9	184	1 YH22_VACCV	HYPOTHETICAL 21.7 KDA	8.75e+00
19	6	5.9	196	1 HTGA_ECOLI	HEAT SHOCK PROTEIN HTG	8.75e+00
20	6	5.9	198	1 VIF_BIV27	VIRION INFECTIVITY FAC	8.75e+00
21	6	5.9	198	1 VIF_BIV06	VIRION INFECTIVITY FAC	8.75e+00
22	6	5.9	218	1 GCHI_HAEIN	GTP CYCLOHYDROLASE I (8.75e+00
23	6	5.9	219	1 PAA4_ECOLI	RESOLVASE.	8.75e+00

24	6	5.9	221	1 GCHI_ECOLI	GTP CYCLOHYDROLASE I (8.75e+00
25	6	5.9	225	1 PDGB_RAT	PLATELET-DERIVED GROWT	8.75e+00
26	6	5.9	226	1 TSIS_SMSAV	PDGF-RELATED TRANSFORM	8.75e+00
27	6	5.9	228	1 FANE_ECOLI	CHAPERONE PROTEIN FANE	8.75e+00
28	6	5.9	231	1 NARW_ECOLI	RESPIRATORY NITRATE RE	8.75e+00
29	6	5.9	241	1 PDGB_SHEEP	PLATELET-DERIVED GROWT	8.75e+00
30	6	5.9	241	1 PDGB_MOUSE	PLATELET-DERIVED GROWT	8.75e+00
31	6	5.9	241	1 PDGB_HUMAN	PLATELET-DERIVED GROWT	8.75e+00
32	6	5.9	245	1 PDGB_FELCA	PLATELET-DERIVED GROWT	8.75e+00
33	6	5.9	248	1 VGL2_EBV	PROBABLE MEMBRANE GLYC	8.75e+00
34	6	5.9	254	1 YKT0_CAEEI	HYPOTHETICAL 29.1 KDA	8.75e+00
35	6	5.9	255	1 RNPH_AQUAE	RIBONUCLEASE PH (EC 2.	8.75e+00
36	6	5.9	264	1 MS2A_DROME	MALE ACCESSORY GLAND S	8.75e+00
37	6	5.9	275	1 YAIM_HAEIN	HYPOTHETICAL PROTEIN H	8.75e+00
38	6	5.9	279	1 HEM3_CHLVI	PORPHOBILINOGEN DEAMIN	8.75e+00
39	6	5.9	288	1 YLW9_CAEEI	HYPOTHETICAL 32.9 KDA	8.75e+00
40	6	5.9	288	1 PYR3_CAEEI	PHYCOBILISOME 39 KDA L	8.75e+00
41	6	5.9	296	1 RS2_RICPR	30S RIBOSOMAL PROTEIN	8.75e+00
42	6	5.9	325	1 SFMH_ECOLI	RFMBH PROTEIN PRECURSOR	8.75e+00
43	6	5.9	326	1 RLUD_BORBU	RIBOSOMAL LARGE SUBUNI	8.75e+00
44	6	5.9	328	1 VGH_BPPHX	MINOR SPIKE PROTEIN (H	8.75e+00
45	6	5.9	328	1 VGH_BPSL3	MINOR SPIKE PROTEIN (H	8.75e+00
46	6	5.9	330	1 VGH_BPAL3	MINOR SPIKE PROTEIN (H	8.75e+00
47	6	5.9	332	1 VGH_BPPHK	MINOR SPIKE PROTEIN (H	8.75e+00
48	6	5.9	337	1 VGH_BFG4	MINOR SPIKE PROTEIN (H	8.75e+00
49	6	5.9	366	1 YG5F_YEAST	PUTATIVE MITOCHONDRIAL	8.75e+00
50	6	5.9	398	1 AAAD_HUMAN	ARYLACETAMIDE DEACETYL	8.75e+00
51	6	5.9	400	1 PAP3_XENLA	POLY(A) POLYMERASE TYP	8.75e+00
52	6	5.9	404	1 CAG5_CHICK	ALPHA-N-ACETYLGALACTOS	8.75e+00
53	6	5.9	407	1 DEOB_ECOLI	PHOSPHOPENTOMUTASE (EC	8.75e+00
54	6	5.9	415	1 GLYA_BACSU	SERINE HYDROXYMETHYLTR	8.75e+00
55	6	5.9	420	1 GLYA_STRCO	SERINE HYDROXYMETHYLTR	8.75e+00
56	6	5.9	425	1 GLA2_MYCTU	SERINE HYDROXYMETHYLTR	8.75e+00
57	6	5.9	426	1 GLA1_MYCTU	SERINE HYDROXYMETHYLTR	8.75e+00
58	6	5.9	426	1 GLYA_MYCLE	SERINE HYDROXYMETHYLTR	8.75e+00
59	6	5.9	427	1 GLYA_SYNY3	SERINE HYDROXYMETHYLTR	8.75e+00
60	6	5.9	431	1 CD44_MESAU	CD44 ANTIGEN PRECURSOR	8.75e+00
61	6	5.9	434	1 MOTC_RHIME	CHEMOTAXIS MOTC PROTEI	8.75e+00
62	6	5.9	434	1 GLYA_METEX	SERINE HYDROXYMETHYLTR	8.75e+00
63	6	5.9	444	1 SLAP_LACAC	S-LAYER PROTEIN PRECUR	8.75e+00
64	6	5.9	447	1 VANS_ENTFA	SENSOR PROTEIN VANSB (8.75e+00
65	6	5.9	468	1 AMYB_BACFI	BETA-AMYLASE PRECURSOR	8.75e+00
66	6	5.9	476	1 COX1_PLACH	CYTOCHROME C OXIDASE P	8.75e+00
67	6	5.9	476	1 COX1_PLABE	CYTOCHROME C OXIDASE P	8.75e+00
68	6	5.9	508	1 MKK1_YEAST	PROTEIN KINASE MKK1/SS	8.75e+00
69	6	5.9	510	1 CD39_HUMAN	VASCULAR ATP-DIPHOSPHO	8.75e+00
70	6	5.9	522	1 GDS1_YEAST	GDS1 PROTEIN.	8.75e+00
71	6	5.9	537	1 UGT5_CAEEI	PUTATIVE UDP-GLUCURONO	8.75e+00
72	6	5.9	542	1 OPPA_SALTY	PERIPLASMIC OLIGOPEPTI	8.75e+00
73	6	5.9	558	1 Y37B_MYCLE	HYPOTHETICAL 61.5 KDA	8.75e+00
74	6	5.9	566	1 HEMA_TAPIL	HEMAGGLUTININ PRECURSO	8.75e+00
75	6	5.9	566	1 HEMA_TAGU2	HEMAGGLUTININ PRECURSO	8.75e+00
76	6	5.9	577	1 THT1_SCHPO	NUCLEAR FUSION PROTEIN	8.75e+00
77	6	5.9	619	1 HEX3_YEAST	HEXOSE METABOLISM-RELA	8.75e+00
78	6	5.9	652	1 PAP_HUMAN	POLY(A) POLYMERASE (EC	8.75e+00
79	6	5.9	655	1 CD44_MOUSE	CD44 ANTIGEN PRECURSOR	8.75e+00
80	6	5.9	688	1 YH10_YEAST	PUTATIVE PROLYL-TRNA S	8.75e+00
81	6	5.9	697	1 Y441_HUMAN	HYPOTHETICAL ZINC FING	8.75e+00
82	6	5.9	712	1 ENV_HV252	ENVELOPE POLYPROTEIN G	8.75e+00
83	6	5.9	713	1 CDG2_PAEMA	CYCLAMALTODEXTRIN GLUC	8.75e+00
84	6	5.9	715	1 PAP1_XENLA	POLY(A) POLYMERASE TYP	8.75e+00
85	6	5.9	738	1 PAP_BOVIN	POLY(A) POLYMERASE (EC	8.75e+00
86	6	5.9	754	1 BGLB_CLOTH	THERMOSTABLE BETA-GLUC	8.75e+00
87	6	5.9	776	1 YLAA_CAEEI	HYPOTHETICAL 86.9 KDA	8.75e+00
88	6	5.9	780	1 Z084_XENLA	OOCYTE ZINC FINGER PRO	8.75e+00
89	6	5.9	810	1 KJUS_YEAST	PROBABLE SERINE/THREON	8.75e+00
90	6	5.9	840	1 YHT1_YEAST	HYPOTHETICAL 95.1 KDA	8.75e+00
91	6	5.9	843	1 MVPA_DICDI	MAJOR VAULT PROTEIN AL	8.75e+00
92	6	5.9	859	1 ENV_HV25T	ENVELOPE POLYPROTEIN G	8.75e+00
93	6	5.9	867	1 DPOL_RICPR	DNA POLYMERASE I (EC 2	8.75e+00
94	6	5.9	874	1 DP3A_MYCGE	DNA POLYMERASE III, AL	8.75e+00
95	6	5.9	960	1 CAP2_MESCR	PHOSPHOENOLPYRUVATE CA	8.75e+00
96	6	5.9	967	1 CAP1_FLAPR	PHOSPHOENOLPYRUVATE CA	8.75e+00

97	6	5.9	990	1	DPOL_NPVCF	DNA POLYMERASE (EC 2.7	8.75e+00	170	5	5.0	159	1	YXBD_BACSU	HYPOTHETICAL 18.1 KDA	4.58e+02
98	6	5.9	1019	1	ENTK_HUMAN	ENTEROPEPTIDASE PRECUR	8.75e+00	171	5	5.0	159	1	GLBW_CHITP	GLOBIN CTT-W PRECURSOR	4.58e+02
99	6	5.9	1120	1	MPD_RICPR	TRANSCRIPTION-REPAIR C	8.75e+00	172	5	5.0	160	1	YRN5_CAEEL	HYPOTHETICAL 16.4 KDA	4.58e+02
100	6	5.9	1196	1	AMVB_PAEPO	BETA/ALPHA-AMYLASE PRE	8.75e+00	173	5	5.0	161	1	RNKD_PONPY	NONSECRETORY RIBONUCLE	4.58e+02
101	6	5.9	1266	1	SVI_HUMAN	ISOLEUCYL-TRNA SYNTHET	8.75e+00	174	5	5.0	162	1	YNI1_METYL	HYPOTHETICAL PROTEIN I	4.58e+02
102	6	5.9	1313	1	VGLM_PTPV	M POLYPROTEIN PRECURSOR	8.75e+00	175	5	5.0	163	1	PTSN_ECOLI	NITROGEN REGULATORY II	4.58e+02
103	6	5.9	1387	1	RISC_RAT	REGULATOR OF G-PROTEIN	8.75e+00	176	5	5.0	165	1	Y316_METJA	HYPOTHETICAL PROTEIN R	4.58e+02
104	6	5.9	1475	1	APU_THET	AMYLOPULLULANASE PRECU	8.75e+00	177	5	5.0	171	1	RECX_MYCLE	REGULATORY PROTEIN R	4.58e+02
105	6	5.9	1505	1	SIMA_DROME	SIMILAR PROTEIN.	8.75e+00	178	5	5.0	173	1	Y013_BPHPL	PUTATIVE ADENINE-SPECI	4.58e+02
106	6	5.9	3707	1	PGBM_MOUSE	BASEMENT MEMBRANE-SPEC	8.75e+00	179	5	5.0	175	1	Y901_MYCTU	HYPOTHETICAL 18.9 KDA	4.58e+02
107	6	5.9	3759	1	TRX_DROME	TRITHORAX PROTEIN.	8.75e+00	180	5	5.0	175	1	DEST_MYXXA	DEVELOPMENT-SPECIFIC P	4.58e+02
108	5	5.0	27	1	K739_PICKL	KILLER TOXIN KT395 (FR	4.58e+02	181	5	5.0	175	1	ATPF_CYACA	ATP SYNTHASE B CHAIN (4.58e+02
109	5	5.0	44	1	RPC3_BPHKO	REGULATORY PROTEIN CII	4.58e+02	182	5	5.0	176	1	COPZ_HUMAN	COATOMER ZETA SUBUNIT	4.58e+02
110	5	5.0	45	1	CSPA_LEIST	MAJOR COLD-SHOCK PROTE	4.58e+02	183	5	5.0	177	1	COPZ_BOVIN	COATOMER ZETA SUBUNIT	4.58e+02
111	5	5.0	53	1	NCE1_YEAST	NON-CLASSICAL EXPORT P	4.58e+02	184	5	5.0	178	1	CASK_RAT	KAPPA CASEIN PRECURSOR	4.58e+02
112	5	5.0	57	1	CSPB_BACGO	COLD SHOCK PROTEIN CSP	4.58e+02	185	5	5.0	179	1	RBS1_FRIAG	RIBULOSE BISPHOSPHATE	4.58e+02
113	5	5.0	68	1	CSPC_SALTY	COLD SHOCK-LIKE PROTEI	4.58e+02	186	5	5.0	179	1	Y366_HAEIN	HYPOTHETICAL PROTEIN H	4.58e+02
114	5	5.0	68	1	CSPC_ECOLI	COLD SHOCK-LIKE PROTEI	4.58e+02	187	5	5.0	183	1	YREP_STRAM	HYPOTHETICAL 19.3 KDA	4.58e+02
115	5	5.0	72	1	YF60_MYCTU	HYPOTHETICAL 8.2 KDA P	4.58e+02	188	5	5.0	183	1	TF2D_METJA	TATA-BOX BINDING PROTE	4.58e+02
116	5	5.0	74	1	IPB2_YEAST	PROTEASE B INHIBITORS	4.58e+02	189	5	5.0	183	1	HGXK_TRIFO	HYPOXANTHINE-GUANINE-X	4.58e+02
117	5	5.0	77	1	YUR2_VACCV	HYPOTHETICAL HOST RANG	4.58e+02	190	5	5.0	185	1	YXAK_BACSU	HYPOTHETICAL 19.5 KDA	4.58e+02
118	5	5.0	85	1	R37A_MYXGL	60S RIBOSOMAL PROTEIN	4.58e+02	191	5	5.0	185	1	EPF_TREPA	ELONGATION FACTOR P (E	4.58e+02
119	5	5.0	87	1	YXGM_ECOLI	HYPOTHETICAL 9.9 KDA P	4.58e+02	192	5	5.0	187	1	YCF4_CYAPA	HYPOTHETICAL 21.2 KDA	4.58e+02
120	5	5.0	89	1	PHPP_ALCEB	PHOSPHOCARRIER PROTEIN	4.58e+02	193	5	5.0	188	1	PAGC_SALTY	VIRULENCE MEMBRANE PRO	4.58e+02
121	5	5.0	93	1	VSPM_IBVP3	PUTATIVE SMALL MEMBRAN	4.58e+02	194	5	5.0	189	1	COAT_TYMYA	COAT PROTEIN (VIRION P	4.58e+02
122	5	5.0	94	1	ESA6_MYCTU	6 KDA EARLY SECRETORY	4.58e+02	195	5	5.0	190	1	YNR4_YEAST	HYPOTHETICAL 21.0 KDA	4.58e+02
123	5	5.0	96	1	SY01_HUMAN	SMALL INDUCIBLE CYTOKI	4.58e+02	196	5	5.0	191	1	Y222_HAEIN	HYPOTHETICAL PROTEIN H	4.58e+02
124	5	5.0	107	1	Y453_METJA	HYPOTHETICAL PROTEIN M	4.58e+02	197	5	5.0	193	1	YXGM_HAEIN	PUTATIVE FERREDOXIN-LI	4.58e+02
125	5	5.0	109	1	N016_SOYBN	MODULIN 16 PRECURSOR (4.58e+02	198	5	5.0	194	1	PABA_BACSU	PARA-AMINOBENZOATE/ANT	4.58e+02
126	5	5.0	110	1	YHIT_CHLPN	HYPOTHETICAL HIT-LIKE	4.58e+02	199	5	5.0	194	1	PTH_SYNY3	IMIDAZOLEGLYCEROL-PHOS	4.58e+02
127	5	5.0	111	1	RNPL_RANCA	RIBONUCLEASE, LIVER (E	4.58e+02	200	5	5.0	195	1	H12_WHEAT	MORPHOGENETIC PROTEIN	4.58e+02
128	5	5.0	112	1	YNB7_YEAST	HYPOTHETICAL 12.5 KDA	4.58e+02	201	5	5.0	195	1	VP17_BPHG	SUPEROXIDE DISMUTASE [4.58e+02
129	5	5.0	113	1	HESB_ANASP	HESB PROTEIN.	4.58e+02	202	5	5.0	198	1	YR71_ECOLI	HYPOTHETICAL 21.9 KDA	4.58e+02
130	5	5.0	114	1	YNI0_MYCTU	HYPOTHETICAL 12.8 KDA	4.58e+02	203	5	5.0	200	1	SODM_GANMI	ATP SYNTHASE DELTA' CH	4.58e+02
131	5	5.0	121	1	Y121_ECOLI	INSERTION ELEMENT IS2	4.58e+02	204	5	5.0	200	1	ATP4_IPOBA	HYPOTHETICAL 22.6 KDA	4.58e+02
132	5	5.0	122	1	Y14K_PMW	HYPOTHETICAL 14.1 KDA	4.58e+02	205	5	5.0	202	1	Y1BF_ECOLI	HOLLIDAY JUNCTION DNA	4.58e+02
133	5	5.0	123	1	RR13_ODOSI	CHLOROPLAST 30S RIBOSO	4.58e+02	206	5	5.0	203	1	RUVA_ECOLI	SARCOSE OXIDASE GAMM	4.58e+02
134	5	5.0	124	1	CDX1_RAT	HOMEOBOX PROTEIN CDX-1	4.58e+02	207	5	5.0	203	1	SOXG_COR51	ALPHA-AMYLASE/SUBTILIS	4.58e+02
135	5	5.0	124	1	YCK3_CHLRE	HYPOTHETICAL 14.4 KDA	4.58e+02	208	5	5.0	203	1	IAAS_HORVU	ALPHA-AMYLASE/SUBTILIS	4.58e+02
136	5	5.0	125	1	YES2_YEAST	HYPOTHETICAL 14.3 KDA	4.58e+02	209	5	5.0	206	1	YXGM_ECOLI	HYPOTHETICAL 22.2 KDA	4.58e+02
137	5	5.0	127	1	GR14_NEOCA	ANTIGEN NC14.1 (FRAGME	4.58e+02	210	5	5.0	211	1	YF78_CAEEL	HYPOTHETICAL 24.3 KDA	4.58e+02
138	5	5.0	128	1	YPRR_ECOLI	HYPOTHETICAL PROTEIN R	4.58e+02	211	5	5.0	213	1	VATO_YEAST	VACUOLAR ATP SYNTHASE	4.58e+02
139	5	5.0	129	1	AZUR_PSEPB	FLAGELLAR BASAL-BODY R	4.58e+02	212	5	5.0	214	1	PITX_RHIME	PUTATIVE PIT ACCESSORY	4.58e+02
140	5	5.0	129	1	FLGB_BORBU	FLAGELLAR BASAL-BODY R	4.58e+02	213	5	5.0	215	1	R56_MYCPN	30S RIBOSOMAL PROTEIN	4.58e+02
141	5	5.0	135	1	YRN6_CAEEL	HYPOTHETICAL 14.2 KDA	4.58e+02	214	5	5.0	216	1	YF22_CAEEL	MINOR FIBRIL SUBUNIT	4.58e+02
142	5	5.0	136	1	YRN6_CAEEL	E6 PROTEIN.	4.58e+02	215	5	5.0	218	1	YF22_CAEEL	HYPOTHETICAL 24.2 KDA	4.58e+02
143	5	5.0	137	1	YF6_BPV1	PHOSPHOLIPASE A2, BASI	4.58e+02	216	5	5.0	218	1	FGFF_MOUSE	FIBROBLAST GROWTH FACT	4.58e+02
144	5	5.0	138	1	PAZB_TRIFL	HYPOTHETICAL 15.0 KDA	4.58e+02	217	5	5.0	219	1	YQBP_BACSU	GLUTATHIONE PEROXIDASE	4.58e+02
145	5	5.0	139	1	YLF5_CAEEL	HEMOGLOBIN ALPHA-1 CHA	4.58e+02	218	5	5.0	222	1	GPWA_PSEWI	SUPEROXIDE DISMUTASE I	4.58e+02
146	5	5.0	141	1	HBA1_XENBO	HEMOGLOBIN ALPHA-2 CHA	4.58e+02	219	5	5.0	223	1	SODM_ONCVO	SUPEROXIDE DISMUTASE I	4.58e+02
147	5	5.0	141	1	HBA2_XENBO	HEMOGLOBIN ALPHA-2 CHA	4.58e+02	220	5	5.0	225	1	YVS7_CAEEL	HYPOTHETICAL 26.0 KDA	4.58e+02
148	5	5.0	143	1	YEN6_YEAST	HYPOTHETICAL PROTEIN H	4.58e+02	221	5	5.0	225	1	YB67_AERPE	HYPOTHETICAL PROTEIN A	4.58e+02
149	5	5.0	143	1	Y1FN_HAEIN	GIANT HEMOGLOBIN AIII	4.58e+02	222	5	5.0	226	1	B10D_MYCTU	DETHIOBIOTIN SYNTHETAS	4.58e+02
150	5	5.0	144	1	GLB3_LAMSP	GIANT HEMOGLOBIN AIII	4.58e+02	223	5	5.0	226	1	B10D_MYCTU	DETHIOBIOTIN SYNTHETAS	4.58e+02
151	5	5.0	144	1	RL22_MYCGE	50S RIBOSOMAL PROTEIN	4.58e+02	224	5	5.0	227	1	YQVA_METTF	HISTIDINE TRANSPORT SY	4.58e+02
152	5	5.0	144	1	SODM_PARGL	SUPEROXIDE DISMUTASE I	4.58e+02	225	5	5.0	228	1	HISQ_ECOLI	HISTIDINE TRANSPORT SY	4.58e+02
153	5	5.0	144	1	PSAH_SPIOL	PHOTOSYSTEM I REACTION	4.58e+02	226	5	5.0	230	1	ALKL_PSEOL	OUTER MEMBRANE PROTEIN	4.58e+02
154	5	5.0	144	1	GLBB_RIPPA	GIANT HEMOGLOBINS B CH	4.58e+02	227	5	5.0	231	1	YHHW_ECOLI	26.3 KDA PROTEIN IN GN	4.58e+02
155	5	5.0	145	1	HS12_CAEEL	HEAT SHOCK PROTEIN HSP	4.58e+02	228	5	5.0	231	1	AQP2_ECOLI	AQUAPORIN 2 (BACTERIAL	4.58e+02
156	5	5.0	146	1	LCRR_YERPS	LOW CALCIUM RESPONSE L	4.58e+02	229	5	5.0	231	1	VIC1_AGRTE	VIRC1 PROTEIN.	4.58e+02
157	5	5.0	146	1	LCRR_YERPS	LOW CALCIUM RESPONSE L	4.58e+02	230	5	5.0	231	1	LUXF_PHOPO	NON-FLUORESCENT FLAVOP	4.58e+02
158	5	5.0	147	1	YV59_CAEEL	HYPOTHETICAL 15.2 KDA	4.58e+02	231	5	5.0	231	1	PRC3_CAEEL	PROBABLE PROTEASOME CO	4.58e+02
159	5	5.0	147	1	CALM_YEAST	CALMODULIN.	4.58e+02	232	5	5.0	236	1	BAC3_HAELV	BACTERIAL RHODOPSIN CS	4.58e+02
160	5	5.0	148	1	AZUR_PSEAE	AZURIN PRECURSOR.	4.58e+02	233	5	5.0	236	1	ECSC_BACSU	PROTEIN ECSC.	4.58e+02
161	5	5.0	149	1	YQ2U_YEAST	HYPOTHETICAL 17.4 KDA	4.58e+02	234	5	5.0	237	1	ATPO_ARATH	ATP SYNTHASE DELTA CHA	4.58e+02
162	5	5.0	151	1	YGS2_ANACE	HYPOTHETICAL PROTEIN I	4.58e+02	235	5	5.0	240	1	Y586_MYCTU	HYPOTHETICAL TRANSCRIP	4.58e+02
163	5	5.0	151	1	CSGA_ECOLI	MAJOR CURLIN SUBUNIT P	4.58e+02	236	5	5.0	241	1	Y293_MYCPN	HYPOTHETICAL PROTEIN M	4.58e+02
164	5	5.0	152	1	GLB3_LUCPE	HEMOGLOBIN III (HB III	4.58e+02	237	5	5.0	241	1	PCXB_ACICA	PROTEOCATECHUATE 3,4-DI	4.58e+02
165	5	5.0	153	1	EMX1_HUMAN	HOMEOBOX PROTEIN EMX1	4.58e+02	238	5	5.0	242	1	REIA_CHLTR	RIBOSE 5-PHOSPHATE ISO	4.58e+02
166	5	5.0	155	1	Y721_AGRVI	HYPOTHETICAL 16.3 KDA	4.58e+02	239	5	5.0	243	1	TRPA_CYAME	TRYPTOPHAN SYNTHASE AL	4.58e+02
167	5	5.0	156	1	YMI9_WHEAT	HYPOTHETICAL 18 KDA PR	4.58e+02	240	5	5.0	245	1	YUBG_ECOLI	HYPOTHETICAL 26.3 KDA	4.58e+02
168	5	5.0	156	1	I3MS_BRANA	MICROSPORE-SPECIFIC PR	4.58e+02	241	5	5.0	245	1	PYHD_NPVLID	POLYHEDRIN (MAJOR OCCL	4.58e+02
169	5	5.0	157	1	RS7_BORBU	30S RIBOSOMAL PROTEIN	4.58e+02	242	5	5.0	246	1	RPSB_MYXXA	RNA POLYMERASE SIGMA-B	4.58e+02

243	5	5.0	1	YK01_MYCTU	HYPOTHETICAL 28.7 KDA	4.58e+02	316	1	CCPB_BACSU	CATABOLITE CONTROL PRO	4.58e+02
244	5	5.0	1	AGL4_ARATH	FLORAL HOMEOIC PROTEI	4.58e+02	317	1	CDX2_MOUSE	HOMEOBOX PROTEIN CDX-2	4.58e+02
245	5	5.0	1	Y545_METJA	HYPOTHETICAL PROTEIN M	4.58e+02	318	1	MIAA_RICPR	TRNA DELTA(2)-ISOPENTE	4.58e+02
246	5	5.0	1	ABCX_ANTSP	PROBABLE ATP-DEPENDENT	4.58e+02	319	5	JTV1_HUMAN	JTV-1 PROTEIN	4.58e+02
247	5	5.0	1	VM11_INMIL	MATRIX PROTEIN M1.	4.58e+02	320	5	ATPG_SNY3	ATP SYNTHASE GAMMA CHA	4.58e+02
248	5	5.0	1	VM11_TAFPW	MATRIX PROTEIN M1.	4.58e+02	321	5	YP76_CAEEL	HYPOTHETICAL 36.7 KDA	4.58e+02
249	5	5.0	1	VM11_TALE1	MATRIX PROTEIN M1.	4.58e+02	322	5	VANB_PSEPU	VANILLATE O-DEMETHYLAS	4.58e+02
250	5	5.0	1	VM11_TAFMR	MATRIX PROTEIN M1.	4.58e+02	323	5	ATPG_ANASP	ATP SYNTHASE GAMMA CHA	4.58e+02
251	5	5.0	1	VM11_TAFOW	MATRIX PROTEIN M1.	4.58e+02	324	5	ATPG_SYNP1	ATP SYNTHASE GAMMA CHA	4.58e+02
252	5	5.0	1	YC88_METJA	HYPOTHETICAL PROTEIN M	4.58e+02	325	5	TRXB_MYCPN	THIOREDOXIN REDUCTASE	4.58e+02
253	5	5.0	1	KD0D_BACSU	2-DEOXY-D-GLUCONATE 3-	4.58e+02	326	5	TRXB_MYCPN	THIOREDOXIN REDUCTASE	4.58e+02
254	5	5.0	1	TRY3_ADAE	TRYPSIN 3A1 PRECURSOR	4.58e+02	327	5	ATPG_SYNP6	ATP SYNTHASE GAMMA CHA	4.58e+02
255	5	5.0	1	GLT1_HAEIN	GLUTAMATE/ASPARTATE TR	4.58e+02	328	5	YZ11_AQUAE	HYPOTHETICAL PROTEIN A	4.58e+02
256	5	5.0	1	YAB0_HAEIN	HYPOTHETICAL AMINO-ACI	4.58e+02	329	5	YIDZ_ECOLI	HYPOTHETICAL TRANSRCU	4.58e+02
257	5	5.0	1	Y224_METJA	HYPOTHETICAL PROTEIN M	4.58e+02	330	5	PRXC_CALFU	CHLOROPEROXIDASE PRECU	4.58e+02
258	5	5.0	1	Y224_METJA	HYPOTHETICAL PROTEIN M	4.58e+02	331	5	CISY_BARVI	CITRATE SYNTHASE (EC 4	4.58e+02
259	5	5.0	1	Y224_METJA	HYPOTHETICAL PROTEIN M	4.58e+02	332	5	VP10_RGDV	NONSTRUCTURAL PROTEIN	4.58e+02
260	5	5.0	1	Y96C_MYCPN	PROBABLE AMINO-ACID AB	4.58e+02	333	5	YCFS_ECOLI	HYPOTHETICAL 34.6 KDA	4.58e+02
261	5	5.0	1	PANB_SNY3	PROBABLE 3-METHYL-2-OX	4.58e+02	334	5	CISY_BARBA	CITRATE SYNTHASE (EC 4	4.58e+02
262	5	5.0	1	AQPB_HUMAN	AQUAPORIN 8.	4.58e+02	335	5	IM30_PEA	CHLOROPLAST MEMBRANE-A	4.58e+02
263	5	5.0	1	ITNB_CHICK	SUCCINATE DEHYDROGENAS	4.58e+02	336	5	CAN2_PIG	CALPAIN 2, LARGE ICATA	4.58e+02
264	5	5.0	1	ITNB_CHICK	INTEGRAL MEMBRANE PROT	4.58e+02	337	5	NORZ_BRAJA	MODULATION PROTEIN 2 (4.58e+02
265	5	5.0	1	COBB_SALTY	PUTATIVE NICOTINATE-NU	4.58e+02	338	5	CYL_HUMAN	CYTOCHROME C1, HEME PR	4.58e+02
266	5	5.0	1	PRC2_HUMAN	PROTEASOME COMPONENT C	4.58e+02	339	5	ASPG_LUPLU	L-ASPARAGINASE (EC 3.5	4.58e+02
267	5	5.0	1	CANS_BOVIN	CALCIUM-DEPENDENT PROT	4.58e+02	340	5	ASPG_LUPAN	L-ASPARAGINASE (EC 3.5	4.58e+02
268	5	5.0	1	YLKP_CAEEL	HYPOTHETICAL 30.3 KDA	4.58e+02	341	5	ASPG_LUPAN	L-ASPARAGINASE (EC 3.5	4.58e+02
269	5	5.0	1	YXEM_BACSU	PROBABLE AMINO-ACID AB	4.58e+02	342	5	YVGR_ECOLI	HYPOTHETICAL 36.0 KDA	4.58e+02
270	5	5.0	1	PC9_DROME	PROTEASOME 29 KDA SUBU	4.58e+02	343	5	MORA_SYPN7	MOLYBDENUM COFACTOR BI	4.58e+02
271	5	5.0	1	OTSB_RHISN	PROBABLE TREHALOSE-PHO	4.58e+02	344	5	COBP_PSESM	COPPER RESISTANCE PROT	4.58e+02
272	5	5.0	1	UBIE_RICPR	PROBABLE TREHALOSE-PHO	4.58e+02	345	5	YH03_YEAST	HYPOTHETICAL 37.4 KDA	4.58e+02
273	5	5.0	1	ITNB_MOUSE	INTEGRAL MEMBRANE PROT	4.58e+02	346	5	IPNS_STRJU	ISOPENICILLIN N SYNTHS	4.58e+02
274	5	5.0	1	YL12_MYCHO	HYPOTHETICAL 31.2 KDA	4.58e+02	347	5	YWC8_BACSU	HYPOTHETICAL 36.6 KDA	4.58e+02
275	5	5.0	1	YC73_HAEIN	HYPOTHETICAL PROTEIN H	4.58e+02	348	5	DPBP_HAEIN	DIPEPTIDE TRANSPORT SY	4.58e+02
276	5	5.0	1	TRPA_PSESY	TRYPTOPHAN SYNTHASE AL	4.58e+02	349	5	RLAO_SULAC	ACIDIC RIBOSOMAL PROTE	4.58e+02
277	5	5.0	1	Y374_MYCPN	HYPOTHETICAL PROTEIN M	4.58e+02	350	5	STSL_ARATH	STRICTOSIDINE SYNTHASE	4.58e+02
278	5	5.0	1	ITNB_BACSU	PROBABLE AMINO-ACID AB	4.58e+02	351	5	TMK8_CAEEL	TMK-8 PROTEIN.	4.58e+02
279	5	5.0	1	HIS1_ARATH	IMIDAZOLEGLYCEROL-PHOS	4.58e+02	352	5	HEM1_CHLTR	GLUTAMYL-TRNA REDUCTAS	4.58e+02
280	5	5.0	1	YM35_MYCTU	HYPOTHETICAL 29.8 KDA	4.58e+02	353	5	GLPX_SHIFL	GLPX PROTEIN.	4.58e+02
281	5	5.0	1	YN97_YEAST	HYPOTHETICAL 31.2 KDA	4.58e+02	354	5	GLPX_ECOLI	GLPX PROTEIN.	4.58e+02
282	5	5.0	1	CYL_RHORI	CYTOCHROME C1 PRECURSO	4.58e+02	355	5	CADH_PETCR	CINNAMYL-ALCOHOL DEHYD	4.58e+02
283	5	5.0	1	VG02_BPT4	TERMINAL DNA PROTECTIN	4.58e+02	356	5	YGBQ_HAEIN	HYPOTHETICAL PROTEIN H	4.58e+02
284	5	5.0	1	TRPC_METTM	INDOLE-3-GLYCEROL PHOS	4.58e+02	357	5	ILVC_RHOMO	KETOL-ACID REDUCTOISOM	4.58e+02
285	5	5.0	1	CAPB_MOUSE	F-ACTIN CAPPING PROTEI	4.58e+02	358	5	TDH_XANCP	THREONINE 3-DEHYDROGEN	4.58e+02
286	5	5.0	1	PRTA_ASPNG	ASPERGILLOPEPSIN II PR	4.58e+02	359	5	MAG_DERFA	ALLERGEN MAG (FRAGMENT	4.58e+02
287	5	5.0	1	CXBL_HUMAN	GAP JUNCTION BETA-1 PR	4.58e+02	360	5	Y535_METJA	HYPOTHETICAL PROTEIN M	4.58e+02
288	5	5.0	1	CXBL_RAT	GAP JUNCTION BETA-1 PR	4.58e+02	361	5	ILVE_HAEIN	BRANCHED-CHAIN AMINO A	4.58e+02
289	5	5.0	1	ALKE_BABBO	ALDO-KETO REDUCTASE (F	4.58e+02	362	5	ASTA_ECOLI	ARGININE N-SUCCINYLTURA	4.58e+02
290	5	5.0	1	ACCD_PORPU	ACETYL-COENZYME A CARB	4.58e+02	363	5	LIMA_PSES5	LIPASE MODULATOR PRECU	4.58e+02
291	5	5.0	1	ILVE_RICPR	PROBABLE BRANCHED-CHAI	4.58e+02	364	5	LIMA_BURCE	LIPASE MODULATOR PRECU	4.58e+02
292	5	5.0	1	MOTV_VIBPA	SODIUM-TYPE FLAGELLAR	4.58e+02	365	5	YJ9S_YEAST	HYPOTHETICAL 39.0 KDA	4.58e+02
293	5	5.0	1	YHAE_ECOLI	HYPOTHETICAL 30.4 KDA	4.58e+02	366	5	NU2M_DIDMA	NADH-UBIQUINONE OXIDOR	4.58e+02
294	5	5.0	1	HK25_CHICK	HOMEOBOX PROTEIN NKX-2	4.58e+02	367	5	TCR_STRRM	TETRACYCLINE RESISTANC	4.58e+02
295	5	5.0	1	YK05_MYCTU	HYPOTHETICAL 30.9 KDA	4.58e+02	368	5	IRF2_HUMAN	INTERFERON REGULATORY	4.58e+02
296	5	5.0	1	SAPR_STRPU	SPERACT PRECURSOR (SPE	4.58e+02	369	5	THI4_ARATH	THIAZOLE BIOSYNTHETIC	4.58e+02
297	5	5.0	1	RN15_YEAST	MRNA 3'-END PROCESSING	4.58e+02	370	5	Y957_METJA	HYPOTHETICAL 38.5 KDA	4.58e+02
298	5	5.0	1	BIEA_HUMAN	BILIVERDIN REDUCTASE A	4.58e+02	371	5	YXAL_BACSU	HYPOTHETICAL 38.5 KDA	4.58e+02
299	5	5.0	1	YHAI_ECOLI	HYPOTHETICAL TRANSRCIP	4.58e+02	372	5	CUP5_GALME	PUPAL CUTICLE PROTEIN	4.58e+02
300	5	5.0	1	PIK3_MOUSE	PUTUTARY HOMEOBOX 3 (4.58e+02	373	5	ALF_XANFL	FRUCTOSE-BISPHOSPHATE	4.58e+02
301	5	5.0	1	PIK3_MOUSE	PUTUTARY HOMEOBOX 3 (4.58e+02	374	5	ALF2_RHOSH	FRUCTOSE-BISPHOSPHATE	4.58e+02
302	5	5.0	1	MYB3_HORV	MYB-RELATED PROTEIN HV	4.58e+02	375	5	KLF2_HUMAN	KRUPPEL-LIKE FACTOR 2	4.58e+02
303	5	5.0	1	PTB_CLOAB	PHOSPHATE BUTYRYLTRANS	4.58e+02	376	5	PMFE_PROMI	PUTATIVE MINOR FIBRIA	4.58e+02
304	5	5.0	1	PIK3_RAT	PUTUTARY HOMEOBOX 3 (4.58e+02	377	5	DHSO_YEAST	SORBITOL DEHYDROGENASE	4.58e+02
305	5	5.0	1	RDGC_ECOLI	RECOMBINATION ASSOCIAT	4.58e+02	378	5	ATPG_CHLRE	ATP SYNTHASE GAMMA CHA	4.58e+02
306	5	5.0	1	PQOB_ACICA	COENZYME PQ SYNTHESIS	4.58e+02	379	5	TGT_RICPR	QUEUINE TRNA-RIBOSYLTR	4.58e+02
307	5	5.0	1	MOV6_BMW	CELL-TO-CELL MOVEMENT	4.58e+02	380	5	YR62_CAEEL	PUTATIVE SERINE/THREON	4.58e+02
308	5	5.0	1	RNH_BPT4	RIBONUCLEASE H (EC 3.1	4.58e+02	381	5	VDH_STRCO	VALINE DEHYDROGENASE (4.58e+02
309	5	5.0	1	Y4CC_RHISN	PROBABLE DNA-INVERTASE	4.58e+02	382	5	YACI_BACSU	HYPOTHETICAL 41.1 KDA	4.58e+02
310	5	5.0	1	YOPD_YEREN	YOPD PROTEIN.	4.58e+02	383	5	ATPG_SPTOL	ATP SYNTHASE GAMMA CHA	4.58e+02
311	5	5.0	1	YOPD_YERPS	YOPD PROTEIN.	4.58e+02	384	5	NADA_SALTY	QUINOLINATE SYNTHETASE	4.58e+02
312	5	5.0	1	ASPG_LUPAR	L-ASPARAGINASE (EC 3.5	4.58e+02	385	5	YMY7_YEAST	HYPOTHETICAL 42.1 KDA	4.58e+02
313	5	5.0	1	ER25_CANAL	C-4 METHYL STEROL OXID	4.58e+02	386	5	AROF_SCHPO	PUTATIVE PHOSPHO-2-DEH	4.58e+02
314	5	5.0	1	YOH1_AZOVI	HYPOTHETICAL 33.2 KDA	4.58e+02	387	5	Y4WE_RHISN	PROBABLE AMINOTRANSFER	4.58e+02
315	5	5.0	1	ATPG_SPTPL	ATP SYNTHASE GAMMA CHA	4.58e+02	388	5	RF2_STRCO	PEPTIDE CHAIN RELEASE	4.58e+02

389	5	5.0	369	1	Y316_MYCPN	HYPOTHETICAL PROTEIN M	4.58e+02	462	5	5.0	428	1	DHE4_SYN3	NADP-SPECIFIC GLUTAMAT	4.58e+02
390	5	5.0	370	1	ATPG_ODOI	ATP SYNTHASE GAMMA CHA	4.58e+02	463	5	5.0	428	1	KINB_BACSU	SEROLATION KINASE B (4.58e+02
391	5	5.0	371	1	YHSC_CLOAB	HYPOTHETICAL 42.4 KDA	4.58e+02	464	5	5.0	429	1	CBXP_ORYSA	SERINE CARBOXYPEPTIDAS	4.58e+02
392	5	5.0	372	1	YKQA_CAEEL	HYPOTHETICAL 38.5 KDA	4.58e+02	465	5	5.0	429	1	CISZ_RHLTR	CITRATE SYNTHASE, PLAS	4.58e+02
393	5	5.0	373	1	3BH1_MESAU	3 BETA-HYDROXYSTEROID	4.58e+02	466	5	5.0	430	1	YNBP_ECOLI	HYPOTHETICAL 49.6 KDA	4.58e+02
394	5	5.0	375	1	TGT_ECOLI	QUEUINE TRNA-RIBOSYLTR	4.58e+02	467	5	5.0	430	1	AMPP_HAEIN	XAA-PRO AMINOPEPTIDASE	4.58e+02
395	5	5.0	375	1	MASB_RAT	MASPIN PRECURSOR (PROT	4.58e+02	468	5	5.0	430	1	CARP_RHIMI	MUCROPEPSIN PRECURSOR	4.58e+02
396	5	5.0	376	1	WEGB_ECOLI	UDP-N-ACETYLGLUCOSAMIN	4.58e+02	469	5	5.0	432	1	ST11_XENLA	SERINE/THREONINE-PROTE	4.58e+02
397	5	5.0	376	1	ACT2_LYTP1	ACTIN, CYTOSKELETAL 2	4.58e+02	470	5	5.0	432	1	GLYA_HVPM	HYPOTHETICAL PROTEIN M	4.58e+02
398	5	5.0	376	1	ACTB_STRPU	ACTIN, CYTOSKELETAL II	4.58e+02	471	5	5.0	434	1	Y181_MYCPN	HYPOTHETICAL PROTEIN M	4.58e+02
399	5	5.0	376	1	ACT1_STRPU	ACTIN, CYTOSKELETAL IB	4.58e+02	472	5	5.0	434	1	GLYA_HVPM	SERINE HYDROXYMETHYLTR	4.58e+02
400	5	5.0	376	1	ACT1_LYTP1	ACTIN, CYTOSKELETAL 1	4.58e+02	473	5	5.0	436	1	THCB_RHOER	CYTOSCHROME P450 116 (E	4.58e+02
401	5	5.0	376	1	OES6_NPVAC	OCCCLUSION-DERIVED VIRU	4.58e+02	474	5	5.0	436	1	GRDB_CLOLI	GLYCINE REDUCTASE COMP	4.58e+02
402	5	5.0	377	1	YD2_DROME	TRANSCRIPTION FACTOR D	4.58e+02	475	5	5.0	440	1	ENOL_CANAL	ENOLASE 1 (EC 4.2.1.11	4.58e+02
403	5	5.0	378	1	Y054_NPVOP	HYPOTHETICAL 42.5 KDA	4.58e+02	476	5	5.0	440	1	COAT_SOCMV	COAT PROTEIN.	4.58e+02
404	5	5.0	378	1	RF2_MYCTU	PEPTIDE CHAIN RELEASE	4.58e+02	477	5	5.0	442	1	WNT_10A	WNT-10A PROTEIN PRECUR	4.58e+02
405	5	5.0	378	1	AGAL_COFAR	ALPHA-GALACTOSIDASE PR	4.58e+02	478	5	5.0	440	1	Y325_HAEIN	HYPOTHETICAL PROTEIN H	4.58e+02
406	5	5.0	379	1	CYB_CANEA	CYTOSCHROME B.	4.58e+02	479	5	5.0	450	1	ENVZ_BACSU	OSMOLARITY SENSOR PROT	4.58e+02
407	5	5.0	379	1	DNBJ_STAAU	DNABJ PROTEIN (HSP40).	4.58e+02	480	5	5.0	451	1	TRPC_HUTCI	TRYPTOPHAN BIOSYNTHESI	4.58e+02
408	5	5.0	380	1	CXAL_BRARE	GAP JUNCTION ALPHA-1 P	4.58e+02	481	5	5.0	451	1	HKX_SCHMA	HEXOKINASE (EC 2.7.1.1	4.58e+02
409	5	5.0	380	1	ENV_STVM2	ENVELOPE POLYPROTEIN G	4.58e+02	482	5	5.0	451	1	ARP2_PLAFA	CLUSTERED-ASPARAGINE-R	4.58e+02
410	5	5.0	380	1	SSUD_ECOLI	ALKANESULFONATE MONOOX	4.58e+02	483	5	5.0	452	1	TRPC_SALTY	TRYPTOPHAN BIOSYNTHESI	4.58e+02
411	5	5.0	380	1	HPPD_STRAW	4-HYDROXYPHENYLPIRUVAT	4.58e+02	484	5	5.0	454	1	ZOX_PHAVU	ZEATIN O-XYLOSYLTRANSF	4.58e+02
412	5	5.0	381	1	YHAD_ECOLI	HYPOTHETICAL 39.1 KDA	4.58e+02	485	5	5.0	455	1	YNU6_YEAST	HYPOTHETICAL 51.6 KDA	4.58e+02
413	5	5.0	382	1	Y022_NPVAC	HYPOTHETICAL 43.8 KDA	4.58e+02	486	5	5.0	455	1	HEML_BACSU	GLUTAMYL-TRNA REDUCTAS	4.58e+02
414	5	5.0	382	1	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	487	5	5.0	456	1	GLMU_HAEIN	UDP-N-ACETYLGLUCOSAMIN	4.58e+02
415	5	5.0	384	1	METK_HAEIN	S-ADENOSYLMETHIONINE S	4.58e+02	488	5	5.0	457	1	CRO_DROME	CROQUEMORT PROTEIN (D-	4.58e+02
416	5	5.0	384	1	POQE_METEX	COENZYME POQ SYNTHESIS	4.58e+02	489	5	5.0	461	1	SEL2_CAEEL	INTEGRAL MEMBRANE PROT	4.58e+02
417	5	5.0	384	1	DFRA_ARATH	DIHYDROFLAVONOL-4-REDU	4.58e+02	490	5	5.0	461	1	PLSB_PHAVU	ARGININE UTILIZATION R	4.58e+02
418	5	5.0	385	1	YHH7_SACDO	HYPOTHETICAL 44.4 KDA	4.58e+02	491	5	5.0	461	1	NIFN_SYNPH	GLYCEROL-3-PHOSPHATE A	4.58e+02
419	5	5.0	386	1	YKD7_CAEEL	HYPOTHETICAL 42.9 KDA	4.58e+02	492	5	5.0	462	1	GSA_MYCTU	GLUTAMATE-1-SEMIALDEHY	4.58e+02
420	5	5.0	386	1	VENV_MCV1	MAJOR ENVELOPE PROTEIN	4.58e+02	493	5	5.0	462	1	FUCO_RAT	TISSUE ALPHA-L-FUCOSID	4.58e+02
421	5	5.0	388	1	VENV_MCV2	MAJOR ENVELOPE PROTEIN	4.58e+02	494	5	5.0	464	1	YK55_CAEEL	HYPOTHETICAL 53.1 KDA	4.58e+02
422	5	5.0	388	1	GALL_LACHE	GALACTOKINASE (EC 2.7.	4.58e+02	495	5	5.0	465	1	YK57_YEAST	HYPOTHETICAL ABC TRANS	4.58e+02
423	5	5.0	392	1	YK33_MYCLE	HYPOTHETICAL 39.3 KDA	4.58e+02	496	5	5.0	465	1	Y065_MYCPN	HYPOTHETICAL 52.0 KDA	4.58e+02
424	5	5.0	392	1	YKH7_YEAST	HYPOTHETICAL 46.0 KDA	4.58e+02	497	5	5.0	466	1	SCRB_SALTY	SUCROSE-6-PHOSPHATE HY	4.58e+02
425	5	5.0	392	1	YQF6_CAEEL	HYPOTHETICAL 45.3 KDA	4.58e+02	498	5	5.0	467	1	HSF3_CHICK	HEAT SHOCK FACTOR PROT	4.58e+02
426	5	5.0	392	1	MTK1_YEAST	HYPOTHETICAL 45.3 KDA	4.58e+02	499	5	5.0	467	1	YAIU_ECOLI	HYPOTHETICAL 50.3 KDA	4.58e+02
427	5	5.0	394	1	MPK1_XENLA	DUAL SPECIFICITY MITOG	4.58e+02	500	5	5.0	467	1	EUTA_ECOLI	ETHANOLAMINE UTILIZATI	4.58e+02
428	5	5.0	394	1	BEKD_HAEIN	CAPSULE POLYSACCHARIDE	4.58e+02	501	5	5.0	470	1	PABB_LACLA	PARA-AMINOBENZONATE SYN	4.58e+02
429	5	5.0	395	1	SYW_SCHPO	PUTATIVE TRYPTOPHANYL-	4.58e+02	502	5	5.0	470	1	ARLY_MYCTU	ARGININOSUCCINATE LYAS	4.58e+02
430	5	5.0	395	1	CARP_SYNRA	SYNCEPHAPEPSIN PRECURS	4.58e+02	503	5	5.0	470	1	RO52_MOUSE	52 KDA RO PROTEIN (SJO	4.58e+02
431	5	5.0	396	1	FLGE_BRUBA	FLAGELLAR HOOK PROTEIN	4.58e+02	504	5	5.0	471	1	Y4A4_SCHPO	HYPOTHETICAL 52.7 KDA	4.58e+02
432	5	5.0	397	1	GLGC_STRCO	GLUCOSE-1-PHOSPHATE AD	4.58e+02	505	5	5.0	473	1	YV33_MYCTU	HYPOTHETICAL 46.9 KDA	4.58e+02
433	5	5.0	397	1	CPXQ_SACER	CYTOSCHROME P450 113A1	4.58e+02	506	5	5.0	473	1	DNB2_ADEAO	EARLY E2A DNA-BINDING	4.58e+02
434	5	5.0	397	1	YQ64_MYCTU	HYPOTHETICAL 42.2 KDA	4.58e+02	507	5	5.0	474	1	GSHB_RAT	GLUTATHIONE SYNTHETASE	4.58e+02
435	5	5.0	400	1	YQ2W_YEAST	HYPOTHETICAL 46.7 KDA	4.58e+02	508	5	5.0	474	1	CREC_ECOLI	SENSOR PROTEIN CREC (E	4.58e+02
436	5	5.0	402	1	OPS4_CANAL	OPAQUE-PHASE-SPECIFIC	4.58e+02	509	5	5.0	475	1	HBDH_ECOLI	PROBABLE 3-HYDROXYBUTY	4.58e+02
437	5	5.0	404	1	PD3C_BRANA	OMEGA-3 FATTY ACID DES	4.58e+02	510	5	5.0	476	1	YHGF_NEIGO	HYPOTHETICAL PROTEIN I	4.58e+02
438	5	5.0	404	1	YXAM_RHISN	HYPOTHETICAL 44.7 KDA	4.58e+02	511	5	5.0	478	1	YSP3_YEAST	SUBTILISIN-LIKE PROTEA	4.58e+02
439	5	5.0	405	1	YFAW_ECOLI	TYROSYL-TRNA SYNTHETAS	4.58e+02	512	5	5.0	478	1	VTNC_MOUSE	VITRONECTIN PRECURSOR	4.58e+02
440	5	5.0	407	1	SYV_MYCGE	PROBABLE RNA HELICASE	4.58e+02	513	5	5.0	479	1	ARLY_CORGL	ARGININOSUCCINATE LYAS	4.58e+02
441	5	5.0	409	1	Y308_MYCPN	PROBABLE RNA HELICASE	4.58e+02	514	5	5.0	479	1	PTSB_VIBAL	PTS SYSTEM, SUCROSE-SP	4.58e+02
442	5	5.0	410	1	YB50_METTW	HYPOTHETICAL PROTEIN M	4.58e+02	515	5	5.0	479	1	APRA_PSEAE	ALKALINE METALLOPROTEI	4.58e+02
443	5	5.0	410	1	YK22_DICDI	TYROSINE-PROTEIN KINAS	4.58e+02	516	5	5.0	483	1	SYE_SYN3	GLUTAMYL-TRNA SYNTHETA	4.58e+02
444	5	5.0	411	1	AGAL_CYATE	ALPHA-GALACTOSIDASE PR	4.58e+02	517	5	5.0	483	1	KG3A_RAT	GLYCOCEN SYNTHASE KINA	4.58e+02
445	5	5.0	412	1	Y070_MYCTU	HYPOTHETICAL 44.9 KDA	4.58e+02	518	5	5.0	484	1	SCRB_VIBAL	GLYCOCEN SYNTHASE KINA	4.58e+02
446	5	5.0	414	1	YFAB_VIHA	3-OXOACYL-[ACYL-CARRIE	4.58e+02	519	5	5.0	484	1	NOT2_MOUSE	SUCROSE-6-PHOSPHATE HY	4.58e+02
447	5	5.0	416	1	GLYA_HELPY	SERINE HYDROXYMETHYLTR	4.58e+02	520	5	5.0	486	1	HS1_HUMAN	MONOCARBOXYLATE TRANSP	4.58e+02
448	5	5.0	417	1	WNIA_MOUSE	WNT-10A PROTEIN PRECUR	4.58e+02	521	5	5.0	486	1	LIPL_CHICK	HEMATOPOIETIC LINEAGE	4.58e+02
449	5	5.0	420	1	HSZ_SYNPH	ATP PHOSPHORIBOSYLTRAN	4.58e+02	522	5	5.0	490	1	CE10_ECOLI	LIPIDOPROTEIN LIPASE PRE	4.58e+02
450	5	5.0	420	1	IL5R_HUMAN	INTERLEUKIN-5 RECEPTOR	4.58e+02	523	5	5.0	491	1	VPN_BPMU	64 KDA VIRION PROTEIN.	4.58e+02
451	5	5.0	420	1	Y181_MYCGE	HYPOTHETICAL PROTEIN M	4.58e+02	524	5	5.0	491	1	PBP_BACSU	PUTATIVE PENICILLIN BI	4.58e+02
452	5	5.0	420	1	KG3B_HUMAN	GLYCOCEN SYNTHASE KINA	4.58e+02	525	5	5.0	491	1	P2C1_CAEEL	PROBABLE PROTEIN PHOSP	4.58e+02
453	5	5.0	420	1	KG3B_RAT	GLYCOCEN SYNTHASE KINA	4.58e+02	526	5	5.0	491	1	CPB2_RAT	CYTOSCHROME P450 2B2 (E	4.58e+02
454	5	5.0	421	1	CHDE_MUCRO	CHITIN DEACETYLASE PRE	4.58e+02	527	5	5.0	492	1	GLPK_AQUAE	GLYCEROL KINASE (EC 2.	4.58e+02
455	5	5.0	424	1	Y12A_ECOLI	INSERTION ELEMENT ISZA	4.58e+02	528	5	5.0	492	1	HUPR_RHOCA	HYDROGENASE TRANSCRIPT	4.58e+02
456	5	5.0	424	1	THC2_METTH	PROBABLE THIAMINE BIOS	4.58e+02	529	5	5.0	492	1	HUPR_RHOCA	HYDROGENASE TRANSCRIPT	4.58e+02
457	5	5.0	425	1	POXN_DROME	PAIRED BOX POX-NEURO P	4.58e+02	530	5	5.0	492	1	PEPD_MOUSE	XAA-PRO DIPEPTIDASE (E	4.58e+02
458	5	5.0	426	1	LH2_RAT	HOMEBOX PROTEIN LH-2.	4.58e+02	531	5	5.0	492	1	PEPD_HUMAN	XAA-PRO DIPEPTIDASE (E	4.58e+02
459	5	5.0	427	1	YKT5_YEAST	HYPOTHETICAL 47.4 KDA	4.58e+02	532	5	5.0	492	1	DYJ2_HUMAN	DYNEIN LIGHT INTERMEDI	4.58e+02
460	5	5.0	428	1	DD1L_YEAST	DNA-DAMAGE INDUCIBLE P	4.58e+02	533	5	5.0	492	1	OPUE_BACSU	OSMOREGULATED PROLINE	4.58e+02
461	5	5.0	428	1	GLYA_AQUAE	SERINE HYDROXYMETHYLTR	4.58e+02	534	5	5.0	493	1	YNL9_YEAST	HYPOTHETICAL 56.5 KDA	4.58e+02

535	5	5.0	493	1	CETP_HUMAN	4.58e+02	608	5	5.0	562	1	YMX8_YEAST	4.58e+02
536	5	5.0	493	1	HXK_PLAFA	4.58e+02	609	5	5.0	562	1	HEMA_TAJAP	4.58e+02
537	5	5.0	493	1	CETP_MACEA	4.58e+02	610	5	5.0	562	1	EXG2_YEAST	4.58e+02
538	5	5.0	494	1	Y663_METJA	4.58e+02	611	5	5.0	564	1	VAS5_VACCC	4.58e+02
539	5	5.0	495	1	UAXA_ECOLI	4.58e+02	612	5	5.0	565	1	HEMA_TAGUA	4.58e+02
540	5	5.0	495	1	HXKG_ASPNC	4.58e+02	613	5	5.0	566	1	HEMA_TAKIE	4.58e+02
541	5	5.0	496	1	KPS5_YEAST	4.58e+02	614	5	5.0	566	1	HEMA_TALEN	4.58e+02
542	5	5.0	497	1	SPLI1_YEAST	4.58e+02	615	5	5.0	566	1	HEMA_TAPOE	4.58e+02
543	5	5.0	497	1	MEK1_YEAST	4.58e+02	616	5	5.0	566	1	NPPE_BACCE	4.58e+02
544	5	5.0	499	1	YP25_METTF	4.58e+02	617	5	5.0	567	1	PGTA_RAT	4.58e+02
545	5	5.0	499	1	MEP2_YEAST	4.58e+02	618	5	5.0	568	1	MET3_ASPTE	4.58e+02
546	5	5.0	500	1	DP51_YEAST	4.58e+02	619	5	5.0	568	1	RS1_RHIME	4.58e+02
547	5	5.0	500	1	YD91_SCHPO	4.58e+02	620	5	5.0	570	1	HEMA_NDVA	4.58e+02
548	5	5.0	500	1	TACT_CLOPE	4.58e+02	621	5	5.0	571	1	HEMA_NDVH3	4.58e+02
549	5	5.0	500	1	UHP8_SALTY	4.58e+02	622	5	5.0	571	1	PEN3_ADE02	4.58e+02
550	5	5.0	500	1	CBP3_ORISA	4.58e+02	623	5	5.0	571	1	PEN3_ADE05	4.58e+02
551	5	5.0	502	1	PUPP_ECOLI	4.58e+02	624	5	5.0	571	1	HEMA_NDVJ	4.58e+02
552	5	5.0	502	1	DCD2_DROME	4.58e+02	625	5	5.0	571	1	HEMA_NDVI	4.58e+02
553	5	5.0	503	1	YPHE_ECOLI	4.58e+02	626	5	5.0	572	1	ZYX_HUMAN	4.58e+02
554	5	5.0	503	1	HSP1_MOUSE	4.58e+02	627	5	5.0	572	1	SYX_HABIN	4.58e+02
555	5	5.0	506	1	YCX2_EUGGR	4.58e+02	628	5	5.0	572	1	MET3_PENCH	4.58e+02
556	5	5.0	506	1	KPY2_YEAST	4.58e+02	629	5	5.0	574	1	VGLF_HRSVR	4.58e+02
557	5	5.0	507	1	IRX3_MOUSE	4.58e+02	630	5	5.0	574	1	COE2_XENLA	4.58e+02
558	5	5.0	508	1	CROC_DROME	4.58e+02	631	5	5.0	575	1	APB3_HUMAN	4.58e+02
559	5	5.0	509	1	AFG1_YEAST	4.58e+02	632	5	5.0	575	1	COE2_MOUSE	4.58e+02
560	5	5.0	510	1	VLI_HPV38	4.58e+02	633	5	5.0	575	1	MLJX_SHEEP	4.58e+02
561	5	5.0	511	1	VGLG_VSVIG	4.58e+02	634	5	5.0	576	1	UN87_CAEEL	4.58e+02
562	5	5.0	511	1	VGLG_VSVJ	4.58e+02	635	5	5.0	577	1	HEMA_NDVH4	4.58e+02
563	5	5.0	514	1	YMP8_YEAST	4.58e+02	636	5	5.0	577	1	HEMA_NDVH	4.58e+02
564	5	5.0	515	1	INR2_HUMAN	4.58e+02	637	5	5.0	579	1	YHVO_YEAST	4.58e+02
565	5	5.0	515	1	PEPB_MCTU	4.58e+02	638	5	5.0	579	1	GPC2_RAT	4.58e+02
566	5	5.0	516	1	ACHD_BOVIN	4.58e+02	639	5	5.0	580	1	SE10_CAEEL	4.58e+02
567	5	5.0	518	1	TRCB_XENLA	4.58e+02	640	5	5.0	581	1	UAPC_EMENI	4.58e+02
568	5	5.0	518	1	LEU1_BUCRP	4.58e+02	641	5	5.0	581	1	AMV1_SCHPO	4.58e+02
569	5	5.0	518	1	ATPA_MYCGA	4.58e+02	642	5	5.0	584	1	COBA_HUMAN	4.58e+02
570	5	5.0	521	1	YDA5_MCTU	4.58e+02	643	5	5.0	584	1	CNA1_DROME	4.58e+02
571	5	5.0	522	1	GLG1_BETVU	4.58e+02	644	5	5.0	586	1	VOID_BPP2	4.58e+02
572	5	5.0	525	1	YLD4_CAEEL	4.58e+02	645	5	5.0	586	1	RRPO_BMYVF	4.58e+02
573	5	5.0	526	1	YHV1_YEAST	4.58e+02	646	5	5.0	586	1	ENV_PMPV	4.58e+02
574	5	5.0	527	1	VP5_EHDV1	4.58e+02	647	5	5.0	586	1	SYR_STRCO	4.58e+02
575	5	5.0	528	1	SERA_MYCTU	4.58e+02	648	5	5.0	586	1	HOLI_YEAST	4.58e+02
576	5	5.0	528	1	TACY_LISIV	4.58e+02	649	5	5.0	587	1	ENV_SRVI	4.58e+02
577	5	5.0	531	1	YDQ2_SCHPO	4.58e+02	650	5	5.0	589	1	RESE_BACSU	4.58e+02
578	5	5.0	531	1	TYD2_PAPSO	4.58e+02	651	5	5.0	590	1	YNG7_YEAST	4.58e+02
579	5	5.0	532	1	ICP0_HSVB	4.58e+02	652	5	5.0	590	1	VG28_HSV11	4.58e+02
580	5	5.0	533	1	INV_DEBOC	4.58e+02	653	5	5.0	591	1	PPOD_LYCES	4.58e+02
581	5	5.0	534	1	YM22_CAEEL	4.58e+02	654	5	5.0	591	1	SYR_BORBU	4.58e+02
582	5	5.0	535	1	YA88_SCHPO	4.58e+02	655	5	5.0	593	1	CDAS_BACSH	4.58e+02
583	5	5.0	536	1	YEN1_SCHPO	4.58e+02	656	5	5.0	593	1	RC03_NEUCR	4.58e+02
584	5	5.0	536	1	SCD2_SCHPO	4.58e+02	657	5	5.0	594	1	IF2P_ARCFU	4.58e+02
585	5	5.0	538	1	PYRG_HELPY	4.58e+02	658	5	5.0	595	1	Y036_HUMAN	4.58e+02
586	5	5.0	540	1	HXTD_YEAST	4.58e+02	659	5	5.0	599	1	TB22_NEIME	4.58e+02
587	5	5.0	542	1	EAT1_BOVIN	4.58e+02	660	5	5.0	601	1	HMDH_CATRO	4.58e+02
588	5	5.0	543	1	PUR6_PICME	4.58e+02	661	5	5.0	604	1	DEDI_YEAST	4.58e+02
589	5	5.0	543	1	EAT1_RAT	4.58e+02	662	5	5.0	604	1	PPO_VITVI	4.58e+02
590	5	5.0	543	1	EAT1_AMBTI	4.58e+02	663	5	5.0	607	1	YD6C_SCHPO	4.58e+02
591	5	5.0	547	1	CH60_KLEPN	4.58e+02	664	5	5.0	609	1	NUSK_HALGR	4.58e+02
592	5	5.0	548	1	PRN_RENSA	4.58e+02	665	5	5.0	613	1	NUSM_POLOR	4.58e+02
593	5	5.0	550	1	CBS_HUMAN	4.58e+02	666	5	5.0	616	1	HEMA_NDVQ	4.58e+02
594	5	5.0	550	1	PTR2_HUMAN	4.58e+02	667	5	5.0	616	1	SRE2_CAEEL	4.58e+02
595	5	5.0	551	1	HR38_DROME	4.58e+02	668	5	5.0	620	1	RRPO_PURVI	4.58e+02
596	5	5.0	553	1	YMAE_YEAST	4.58e+02	669	5	5.0	621	1	NTTA_HUMAN	4.58e+02
597	5	5.0	553	1	YDIE_ECOLI	4.58e+02	670	5	5.0	621	1	NTTA_MOUSE	4.58e+02
598	5	5.0	554	1	MCRA_METVA	4.58e+02	671	5	5.0	622	1	YF54_METJA	4.58e+02
599	5	5.0	555	1	VGA_BPG4	4.58e+02	672	5	5.0	624	1	YGL0_YEAST	4.58e+02
600	5	5.0	555	1	MIS_MOUSE	4.58e+02	673	5	5.0	625	1	AMYG_NEUCR	4.58e+02
601	5	5.0	556	1	FTHS_CLOCY	4.58e+02	674	5	5.0	625	1	TRF5_YEAST	4.58e+02
602	5	5.0	557	1	G6P1_KIULA	4.58e+02	675	5	5.0	627	1	YHAB_YEAST	4.58e+02
603	5	5.0	557	1	G6P1_MOUSE	4.58e+02	676	5	5.0				
604	5	5.0	558	1	LCB1_YEAST	4.58e+02	677	5	5.0				
605	5	5.0	560	1	TF56_SULSH	4.58e+02	678	5	5.0				
606	5	5.0	560	1	NDOD_AZOCA	4.58e+02	679	5	5.0				
607	5	5.0	561	1	MTV1_BACST	4.58e+02	680	5	5.0				

681	1	RA21_SCHPO	DOUBLE-STRAND-BREAK RE	4.58e+02	754	1	MUTB_MYCTU	PROBABLE METHYLMALONYL	4.58e+02
682	1	YD13_SCHPO	HYPOTHETICAL 70.6 KDA	4.58e+02	755	1	RRE1_HUMAN	RAS-RESPONSIVE ELEMENT	4.58e+02
683	1	Y242_MYCGE	HYPOTHETICAL PROTEIN M	4.58e+02	756	1	YHGF_NEIME	HYPOTHETICAL 83.1 KDA	4.58e+02
684	1	VLG2_SYNV	SPIKE GLYCOPROTEIN PRE	4.58e+02	757	1	RRE1_TAWIS	RNA-DIRECTED RNA POLYM	4.58e+02
685	1	Y242_MYCPN	HYPOTHETICAL PROTEIN M	4.58e+02	758	1	YK4_YEAST	HYPOTHETICAL 87.9 KDA	4.58e+02
686	1	YK3_EUGGR	HYPOTHETICAL 78.0 KDA	4.58e+02	759	1	IF39_YEAST	EUKARYOTIC TRANSLATION	4.58e+02
687	1	YK3_YEAST	HYPOTHETICAL 70.6 KDA	4.58e+02	760	1	YHGF_ECOLI	85.1 KDA PROTEIN IN GR	4.58e+02
688	1	YK3_YEAST	HYPOTHETICAL 70.6 KDA	4.58e+02	761	1	YHGF_ECOLI	HYPOTHETICAL 87.1 KDA	4.58e+02
689	1	YK3_YEAST	PROBABLE ATP-DEPENDENT	4.58e+02	762	1	YMA6_YEAST	OUTER CAPSID PROTEIN V	4.58e+02
690	1	GSH1_HUMAN	GLUTAMATE--CYSTEINE LI	4.58e+02	763	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
691	1	PHC3_RHET	POLY-BETA-HYDROXYBUTYR	4.58e+02	764	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
692	1	YHES_HAEIN	HYPOTHETICAL ABC TRANS	4.58e+02	765	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
693	1	GHR_RABIT	GROWTH HORMONE RECEPTO	4.58e+02	766	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
694	1	GHR_RABIT	GROWTH HORMONE RECEPTO	4.58e+02	767	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
695	1	NO52_RHIME	NITROUS-OXIDE REDUCTAS	4.58e+02	768	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
696	1	PP2A_SPIOI	POLYPHENOL OXIDASE PRE	4.58e+02	769	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
697	1	HRPK_PSESY	PATHOGENICITY LOCUS PR	4.58e+02	770	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
698	1	YK3_YEAST	HYPOTHETICAL 73.6 KDA	4.58e+02	771	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
699	1	VE1_HPV57	REPLICATION PROTEIN E1	4.58e+02	772	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
700	1	NUS4_ASTPE	NADH-UBIQUINONE OXIDOR	4.58e+02	773	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
701	1	VE1_HPV2A	REPLICATION PROTEIN E1	4.58e+02	774	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
702	1	GHR_MOUSE	HIGH MOLECULAR WEIGHT	4.58e+02	775	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
703	1	DMR9_MOUSE	DMR-N9 PROTEIN	4.58e+02	776	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
704	1	SEC9_YEAST	PROTEIN TRANSPORT PROT	4.58e+02	777	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
705	1	YNE7_YEAST	HYPOTHETICAL 74.8 KDA	4.58e+02	778	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
706	1	YK16_YEAST	HYPOTHETICAL 74.7 KDA	4.58e+02	779	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
707	1	ORAT2_RAT	SODIUM-INDEPENDENT ORG	4.58e+02	780	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
708	1	YK16_YEAST	HYPOTHETICAL 76.7 KDA	4.58e+02	781	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
709	1	ERF2_SCHPO	EUKARYOTIC PEPTIDE CHA	4.58e+02	782	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
710	1	PRTP_HSVB2	PROBABLE PROCESSING AN	4.58e+02	783	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
711	1	TKT_HAEIN	TRANSKETOLASE (EC 2.2.	4.58e+02	784	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
712	1	HKG_COMXP	HOST RANGE PROTEIN	4.58e+02	785	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
713	1	NRTC_SYNY3	NITRATE TRANSPORT ATP-	4.58e+02	786	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
714	1	CHEA_THENA	CHEMOTAXIS PROTEIN CHE	4.58e+02	787	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
715	1	KPCA_RABIT	PROTEIN KINASE C, ALPH	4.58e+02	788	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
716	1	KPCA_RAT	PROTEIN KINASE C, ALPH	4.58e+02	789	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
717	1	OSM3_CAEEL	KINESIN-LIKE PROTEIN O	4.58e+02	790	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
718	1	KPCA_HUMAN	PROTEIN KINASE C, ALPH	4.58e+02	791	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
719	1	KPCA_BOVIN	PROTEIN KINASE C, ALPH	4.58e+02	792	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
720	1	Y552_HUMAN	HYPOTHETICAL 79.2 KDA	4.58e+02	793	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
721	1	YADD_SCHPO	HYPOTHETICAL 74.2 KDA	4.58e+02	794	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
722	1	BGH3_HUMAN	TRANSFORMING GROWTH FA	4.58e+02	795	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
723	1	LIPE_AERYH	EXTRACELLULAR LIPASE P	4.58e+02	796	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
724	1	ILVB_YEAST	ACETOLACTATE LIPASE P	4.58e+02	797	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
725	1	YK06_YEAST	CHLORIDE CHANNEL PROTE	4.58e+02	798	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
726	1	YK06_YEAST	HYPOTHETICAL 79.2 KDA	4.58e+02	799	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
727	1	K122_STRPU	KINESIN-II 85 KDA SUBU	4.58e+02	800	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
728	1	CAN2_CHICK	CALPAIN 2, LARGE [CATA	4.58e+02	801	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
729	1	PALL_ORYSA	PHENYLALANINE AMMONIA-	4.58e+02	802	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
730	1	KE3A_HUMAN	KINESIN-LIKE PROTEIN K	4.58e+02	803	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
731	1	PCCA_HUMAN	PROPIONYL-COA CARBOXYL	4.58e+02	804	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
732	1	YK70_YEAST	HYPOTHETICAL 79.4 KDA	4.58e+02	805	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
733	1	RED1_RAT	DOUBLE-STRANDED RNA-SP	4.58e+02	806	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
734	1	PQ22_PORTA	RETROVIRUS-RELATED POL	4.58e+02	807	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
735	1	MUTB_RHIME	METHYLMALONYL-COA MUTA	4.58e+02	808	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
736	1	SSK1_YEAST	OSOMOLARITY TWO-COMPO	4.58e+02	809	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
737	1	YK38_YEAST	HYPOTHETICAL 80.2 KDA	4.58e+02	810	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
738	1	CLPB_MYCGE	CLPB PROTEIN	4.58e+02	811	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
739	1	NCPR_CATRO	NADPH-CYTOCHROME P450	4.58e+02	812	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
740	1	CLPB_MYCPN	CLPB PROTEIN	4.58e+02	813	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
741	1	YK6_CAEEL	HYPOTHETICAL 81.4 KDA	4.58e+02	814	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
742	1	YK76_YEAST	HYPOTHETICAL 82.0 KDA	4.58e+02	815	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
743	1	GUNG_CLOCE	ENDOGLUCANASE G PRECUR	4.58e+02	816	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
744	1	AGAL_YEAST	A-AGGLUTININ ATTACHMEN	4.58e+02	817	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
745	1	IF39_SCHPO	PROBABLE EUKARYOTIC TR	4.58e+02	818	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
746	1	TREB_NEUCR	NEUTRAL TREHALASE (EC	4.58e+02	819	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
747	1	YFJ2_YEAST	HYPOTHETICAL 79.7 KDA	4.58e+02	820	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
748	1	MUTB_STRCM	METHYLMALONYL-COA MUTA	4.58e+02	821	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
749	1	K121_STRPU	KINESIN-II 95 KDA SUBU	4.58e+02	822	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
750	1	KM65_YEAST	PROBABLE SERINE/THREON	4.58e+02	823	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
751	1	AMSE_ERWAM	AMYLOVORAN BIOSYNTHESI	4.58e+02	824	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
752	1	KF3B_MOUSE	KINESIN-LIKE PROTEIN K	4.58e+02	825	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
753	1	TREB_EMENT	NEUTRAL TREHALASE (EC	4.58e+02	826	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02

827	5.0	892	1	TRAL_STRFR	TRANSPOSASE FOR TRANSP	4.58e+02	900	5	5.0	1134	1	YML7_YEAST	HYPOTHETICAL 126.1 KDA	4.58e+02
828	5.0	899	1	YML3_YEAST	HYPOTHETICAL 103.0 KDA	4.58e+02	901	5	5.0	1147	1	NRDC_HUMAN	NARDILYSIN PRECURSOR (4.58e+02
829	5.0	900	1	GUNH_CLOTH	ENDOGLUCANASE H PRECUR	4.58e+02	902	5	5.0	1152	1	YBF7_CAEEL	PROBABLY CATION-TRANSP	4.58e+02
830	5.0	900	1	XPC_MOUSE	DNA-REPAIR PROTEIN COM	4.58e+02	903	5	5.0	1158	1	RI14_HUMAN	NUCLEAR FACTOR R1P140	4.58e+02
831	5.0	901	1	MALT_ECOLI	MALT REGULATORY PROTEI	4.58e+02	904	5	5.0	1159	1	YQ4_CAEEL	HYPOTHETICAL 127.4 KDA	4.58e+02
832	5.0	903	1	YB56_METJA	CELL DIVISION CYCLE PR	4.58e+02	905	5	5.0	1161	1	NRDC_RAT	NARDILYSIN PRECURSOR (4.58e+02
833	5.0	906	1	GLRI_HUMAN	GLUTAMATE RECEPTOR 1 P	4.58e+02	906	5	5.0	1166	1	ADDB_BACSU	ATP-DEPENDENT NUCLEASE	4.58e+02
834	5.0	907	1	NUOG_SALTY	NADH DEHYDROGENASE I C	4.58e+02	907	5	5.0	1169	1	SUV3_DROME	SUPPRESSOR OF VARIEGAT	4.58e+02
835	5.0	907	1	GLR1_RAT	GLUTAMATE RECEPTOR 1 P	4.58e+02	908	5	5.0	1176	1	SLAP_BACSH	SURFACE-LAYER 125 KDA	4.58e+02
836	5.0	907	1	IF42_YEAST	EUKARYOTIC INITIATION	4.58e+02	909	5	5.0	1180	1	ITAL_RAT	INTEGRIN ALPHA-1 PRECU	4.58e+02
837	5.0	918	1	YK62_CAEEL	HYPOTHETICAL 105.4 KDA	4.58e+02	910	5	5.0	1192	1	METH_MYCTU	5-METHYLTETRAHYDROFOLA	4.58e+02
838	5.0	918	1	HXK1_RAT	HEXOKINASE, TYPE I (EC	4.58e+02	911	5	5.0	1194	1	APAF_HUMAN	APOPTOTIC PROTEASE ACT	4.58e+02
839	5.0	920	1	HXK1_MOUSE	HEXOKINASE, TYPE I (EC	4.58e+02	912	5	5.0	1199	1	RPOC_BACSU	DNA-DIRECTED RNA POLYM	4.58e+02
840	5.0	920	1	PMAL_NEUCR	PLASMA MEMBRANE ATPASE	4.58e+02	913	5	5.0	1202	1	NO33_HUMAN	NITRIC-OXIDE SYNTHASE,	4.58e+02
841	5.0	920	1	NIA_CICIN	NITRATE REDUCTASE (EC	4.58e+02	914	5	5.0	1217	1	SVF_FUGRU	VALYL-TRNA SYNTHETASE	4.58e+02
842	5.0	921	1	CN2A_BOVIN	CGMP-DEPENDENT 3',5'-C	4.58e+02	915	5	5.0	1226	1	YMA1_CAEEL	PROBABLE INTEGRIN ALPH	4.58e+02
843	5.0	923	1	NRP_MOUSE	NEUROFILIN PRECURSOR (4.58e+02	916	5	5.0	1238	1	YQ09_CAEEL	HYPOTHETICAL 141.2 KDA	4.58e+02
844	5.0	923	1	NRP_HUMAN	NEUROFILIN PRECURSOR (4.58e+02	917	5	5.0	1243	1	VG37_BPK3	TAIL FIBER PROTEIN GP3	4.58e+02
845	5.0	928	1	CN2A_RAT	CGMP-DEPENDENT 3',5'-C	4.58e+02	918	5	5.0	1250	1	BXE_CLOBO	BOTULINUM NEUROTOXIN T	4.58e+02
846	5.0	938	1	YW91_CAEEL	HYPOTHETICAL 105.4 KDA	4.58e+02	919	5	5.0	1250	1	BXE_CLOBO	BOTULINUM NEUROTOXIN T	4.58e+02
847	5.0	941	1	CN2A_HUMAN	CGMP-DEPENDENT 3',5'-C	4.58e+02	920	5	5.0	1251	1	PIP5_HUMAN	1-PHOSPHATIDYLINOSITOL	4.58e+02
848	5.0	943	1	YLM5_CAEEL	HYPOTHETICAL 105.9 KDA	4.58e+02	921	5	5.0	1257	1	FLIH_CAEEL	FLIGHTLESS-1 PROTEIN H	4.58e+02
849	5.0	952	1	YH19_RHOCA	HYPOTHETICAL 104.1 KDA	4.58e+02	922	5	5.0	1259	1	YTFN_ECOLI	HYPOTHETICAL 136.8 KDA	4.58e+02
850	5.0	953	1	LYAG_MOUSE	LYSOSOMAL ALPHA-GLUCOS	4.58e+02	923	5	5.0	1291	1	VAC4_HELPY	VACUOLATING CYTOTOXIN	4.58e+02
851	5.0	953	1	YC07_YEAST	HYPOTHETICAL 107.9 KDA	4.58e+02	924	5	5.0	1308	1	ERB4_HUMAN	ERBB-4 RECEPTOR PROTEI	4.58e+02
852	5.0	954	1	YEP3_YEAST	HYPOTHETICAL 106.7 KDA	4.58e+02	925	5	5.0	1314	1	TETX_CLOTE	TETANUS TOXIN PRECURSO	4.58e+02
853	5.0	956	1	YEP3_YEAST	HYPOTHETICAL 106.1 KDA	4.58e+02	926	5	5.0	1318	1	VIVD_BPT7	INTERNAL VIRION PROTEI	4.58e+02
854	5.0	958	1	YK71_YEAST	HYPOTHETICAL 108.2 KDA	4.58e+02	927	5	5.0	1331	1	MANB_CALSA	BETA-MANNANASE/ENDOGLU	4.58e+02
855	5.0	959	1	BV21_BTV13	OUTER CAPSID PROTEIN V	4.58e+02	928	5	5.0	1360	1	GLI1_XENLA	ZINC FINGER PROTEIN GL	4.58e+02
856	5.0	959	1	VP2_BTV3V	OUTER CAPSID PROTEIN V	4.58e+02	929	5	5.0	1379	1	M3K5_MOUSE	MITOGEN-ACTIVATED PROT	4.58e+02
857	5.0	960	1	CHS3_NEUCR	CHITIN SYNTHASE 3 (EC	4.58e+02	930	5	5.0	1394	1	IT2A_DROME	POSITION-SPECIFIC ANTI	4.58e+02
858	5.0	964	1	OMPE_CHLTR	PUTATIVE OUTER MEMBRAN	4.58e+02	931	5	5.0	1409	1	HAP1_HAEIN	ADHESION AND PENETRATI	4.58e+02
859	5.0	964	1	IF3A_YEAST	EUKARYOTIC TRANSLATION	4.58e+02	932	5	5.0	1427	1	REST_HUMAN	RESTIN (CYTOPLASMIC LI	4.58e+02
860	5.0	967	1	SYA_BOMMO	ALANYL-TRNA SYNTHETASE	4.58e+02	933	5	5.0	1431	1	DAPK_HUMAN	DEATH-ASSOCIATED PROTE	4.58e+02
861	5.0	971	1	LONN_CAEEL	MITOCHONDRIAL LON PROT	4.58e+02	934	5	5.0	1452	1	VGL2_FIPV	E2 GLYCOPROTEIN PRECUR	4.58e+02
862	5.0	973	1	COPB_YEAST	COATOMER BETA SUBUNIT	4.58e+02	935	5	5.0	1458	1	PHLX_RABIT	PHOSPHOLIPASE ADPRAB-B	4.58e+02
863	5.0	985	1	AGLU_ASPNG	ALPHA-GLUCOSIDASE PREC	4.58e+02	936	5	5.0	1466	1	SP2A_YEAST	SPA2 PROTEIN.	4.58e+02
864	5.0	992	1	UVRA_MICLU	EXCINUCLEASE ABC SUBUN	4.58e+02	937	5	5.0	1481	1	APU_THEET	AMYLOPOLYLULANASE PRECU	4.58e+02
865	5.0	993	1	NISB_MACLA	NISIN BIOSYNTHESIS PRO	4.58e+02	938	5	5.0	1489	1	YGP0_YEAST	HYPOTHETICAL 171.5 KDA	4.58e+02
866	5.0	997	1	DP01_TREPA	DNA POLYMERASE I (EC 2	4.58e+02	939	5	5.0	1502	1	N170_YEAST	NUCLEOPORIN NUP170 (NU	4.58e+02
867	5.0	1002	1	YEMA_DROME	YEMANUCLEIN-ALPHA	4.58e+02	940	5	5.0	1520	1	GLTB_BACSU	GLUTAMATE SYNTHASE [NA	4.58e+02
868	5.0	1003	1	SYA_ARATH	ALANYL-TRNA SYNTHETASE	4.58e+02	941	5	5.0	1547	1	TOP2_BOMMO	DNA TOPOISOMERASE II (4.58e+02
869	5.0	1014	1	CIN1_YEAST	CINI PROTEIN.	4.58e+02	942	5	5.0	1548	1	UGGG_DROME	UDP-GLUCOSE:GLYCOPROTE	4.58e+02
870	5.0	1024	1	HLXA_ECOLI	HEMOLYSIN, PLASMID.	4.58e+02	943	5	5.0	1549	1	YJW2_YEAST	PUTATIVE MEMBRANE GLYC	4.58e+02
871	5.0	1026	1	PTP1_CAEEL	PROTEIN-TYROSINE PHOSP	4.58e+02	944	5	5.0	1549	1	YIR3_YEAST	PUTATIVE MEMBRANE GLYC	4.58e+02
872	5.0	1029	1	END1_YEAST	VACUOLAR BIOGENESIS PR	4.58e+02	945	5	5.0	1581	1	YGLP_BEV	PEPOMER GLYCOPROTEIN	4.58e+02
873	5.0	1035	1	MT10_YEAST	SULFITE REDUCTASE [NAD	4.58e+02	946	5	5.0	1592	1	YH5_YEAST	PROBABLE ATP-DEPENDENT	4.58e+02
874	5.0	1036	1	YQ86_CAEEL	HYPOTHETICAL 118.2 KDA	4.58e+02	947	5	5.0	1595	1	SOS_DROME	SON OF SEVENLESS PROTE	4.58e+02
875	5.0	1039	1	YR71_CAEEL	HYPOTHETICAL 118.2 KDA	4.58e+02	948	5	5.0	1626	1	TP2B_HUMAN	DNA TOPOISOMERASE II,	4.58e+02
876	5.0	1046	1	RPOC_WEIHE	ENDOGLUCANASE/EXOGLUCA	4.58e+02	949	5	5.0	1643	1	RRPO_NMV	RNA REPLICATION PROTEI	4.58e+02
877	5.0	1050	1	Y032_HUMAN	DNA-DIRECTED RNA POLYM	4.58e+02	950	5	5.0	1651	1	VIT6_CAEEL	VITELLOGENIN A2 PRECURS	4.58e+02
878	5.0	1052	1	RPOC_BACAN	HYPOTHETICAL PROTEIN K	4.58e+02	951	5	5.0	1707	1	194K_TRVSY	POTENTIAL 194 KDA PROT	4.58e+02
879	5.0	1053	1	RPOC_LISMU	DNA-DIRECTED RNA POLYM	4.58e+02	952	5	5.0	1743	1	TAGC_DICDI	PRESTALK-SPECIFIC PROT	4.58e+02
880	5.0	1053	1	RPOC_LISIN	DNA-DIRECTED RNA POLYM	4.58e+02	953	5	5.0	1746	1	TENA_PIG	TENASCIN PRECURSOR (TN	4.58e+02
881	5.0	1054	1	RPOC_WEIPA	DNA-DIRECTED RNA POLYM	4.58e+02	954	5	5.0	1763	1	PKSM_BACSU	PUTATIVE POLYKETIDE BI	4.58e+02
882	5.0	1057	1	VP2_AHSV3	OUTER CAPSID PROTEIN V	4.58e+02	955	5	5.0	1807	1	VTA2_XENLA	VITELLOGENIN A2 PRECUR	4.58e+02
883	5.0	1060	1	EG51_XENLA	KINESIN-RELATED MOTOR	4.58e+02	956	5	5.0	1809	1	TSC2_RAT	TUBERIN (TUBEROUS SCLE	4.58e+02
884	5.0	1071	1	YB22_SCHPO	HYPOTHETICAL 123.7 KDA	4.58e+02	957	5	5.0	1868	1	YH00_YEAST	HYPOTHETICAL 210.4 KDA	4.58e+02
885	5.0	1071	1	YB22_SCHPO	HYPOTHETICAL 123.7 KDA	4.58e+02	958	5	5.0	1958	1	UBR1_SCHPO	PROBABLE N-END-RECOGNI	4.58e+02
886	5.0	1077	1	HLBS_DROME	HAIRLESS PROTEIN.	4.58e+02	959	5	5.0	1986	1	WA_EMENT	CONIDIAL GREEN PIGMENT	4.58e+02
887	5.0	1077	1	HLBS_DROME	HAIRLESS PROTEIN.	4.58e+02	960	5	5.0	2029	1	CIN5_HUMAN	SODIUM CHANNEL PROTEIN	4.58e+02
888	5.0	1080	1	MYI_TREPA	SERINE/THREONINE-PROTE	4.58e+02	961	5	5.0	2105	1	LAR_DROME	PROTEIN-TYROSINE PHOSP	4.58e+02
889	5.0	1091	1	AF17_HUMAN	ISOLEUCYL-TRNA SYNTHET	4.58e+02	962	5	5.0	2105	1	POLR_ASGVP	GENOME POLYPROTEIN [CO	4.58e+02
890	5.0	1093	1	AF17_HUMAN	AF-17 PROTEIN.	4.58e+02	963	5	5.0	2110	1	MCAS_MYCBO	MYCOCEROSIC ACID SYNTH	4.58e+02
891	5.0	1101	1	KOFS_YEAST	PROBABLE SERINE/THREON	4.58e+02	964	5	5.0	2136	1	YCF2_MARPO	HYPOTHETICAL 259 KDA P	4.58e+02
892	5.0	1102	1	YMG6_YEAST	HYPOTHETICAL 128.1 KDA	4.58e+02	965	5	5.0	2150	1	SDC3_CAEEL	SDC-3 PROTEIN.	4.58e+02
893	5.0	1107	1	POL2_RRVVS	RNA2 POLYPROTEIN [CONT	4.58e+02	966	5	5.0	2151	1	RRPL_SBOUB	RNA-DIRECTED RNA POLYM	4.58e+02
894	5.0	1114	1	RH18_YEAST	DNA REPAIR PROTEIN RHC	4.58e+02	967	5	5.0	2204	1	RRPL_NDV8	RNA POLYMERASE BETA SU	4.58e+02
895	5.0	1115	1	IRE1_YEAST	SERINE/THREONINE-PROTE	4.58e+02	968	5	5.0	2209	1	Y166_HUMAN	HYPOTHETICAL PROTEIN K	4.58e+02
896	5.0	1117	1	CY74_NEUCR	MITOCHONDRIAL PROTEIN	4.58e+02	969	5	5.0	2211	1	PA5_BOVIN	COAGULATION FACTOR V P	4.58e+02
897	5.0	1124	1	POL_FIVSD	POL POLYPROTEIN [CONTA	4.58e+02	970	5	5.0	2225	1	PYR1_MESAU	CAD PROTEIN [INCLUDES:	4.58e+02
898	5.0	1124	1	POL_FIVPE	POL POLYPROTEIN [CONTA	4.58e+02	971	5	5.0	2329	1	Y589_CAEEL	HYPOTHETICAL 254.3 KDA	4.58e+02
899	5.0	1132	1	BAT3_HUMAN	LARGE PROLINE-RICH PRO	4.58e+02	972	5	5.0	2444	1	NTC1_HUMAN	NEUROGENIC LOCUS NOTCH	4.58e+02

973 5 5.0 2524 1 NOTC_XENLA NEUROGENIC LOCUS NOTCH 4.58e+02
974 5.0 2531 1 NTCL_MOUSE NEUROGENIC LOCUS NOTCH 4.58e+02
975 5.0 2531 1 NTCL_MOUSE NEUROGENIC LOCUS NOTCH 4.58e+02
976 5.0 2628 1 HAGA_PORGI HEMAGGLUTININ A PRECUR 4.58e+02
977 5.0 2688 1 ZEP1_MOUSE ZINC FINGER PROTEIN 40 4.58e+02
978 5.0 2703 1 NOTC_DROME NEUROGENIC LOCUS NOTCH 4.58e+02
979 5.0 2717 1 ZEP1_HUMAN ZINC FINGER PROTEIN 40 4.58e+02
980 5.0 2748 1 NUM1_YEAST NUCLEAR MIGRATION PROT 4.58e+02
981 5.0 2871 1 FBNI_BOVIN FIBRILLIN 1 PRECURSOR 4.58e+02
982 5.0 2871 1 FBNI_BOVIN FIBRILLIN 1 PRECURSOR 4.58e+02
983 5.0 3005 1 ZH2_DROME ZINC-FINGER PROTEIN 2 4.58e+02
984 5.0 3092 1 IRAL_YEAST INHIBITORY REGULATOR P 4.58e+02
985 5.0 3140 1 POLG_PPVRA GENOME POLYPROTEIN [CO 4.58e+02
986 5.0 3140 1 POLG_PPVSK GENOME POLYPROTEIN [CO 4.58e+02
987 5.0 3141 1 POLG_PPVD GENOME POLYPROTEIN [CO 4.58e+02
988 5.0 3144 1 VP13_YEAST VACUOLAR PROTEIN SORTI 4.58e+02
989 5.0 3206 1 POLG_PSBMV GENOME POLYPROTEIN [CO 4.58e+02
990 5.0 3418 1 BRC2_HUMAN BREAST CANCER TYPE 2 S 4.58e+02
991 5.0 3433 1 UTRN_HUMAN UTRAPHIN (DYSTROPHIN-R 4.58e+02
992 5.0 3587 1 SRFL_BACSU SURFACTIN SYNTHETASE S 4.58e+02
993 5.0 3744 1 YHP9_YEAST HYPOTHETICAL 433.2 KDA 4.58e+02
994 5.0 3924 1 ANK2_HUMAN ANKYRIN 2 (BRAIN ANKYR 4.58e+02
995 5.0 4349 1 DYHC_FUSSO DYNEIN HEAVY CHAIN, CY 4.58e+02
996 5.0 4466 1 DYHC_ANTCR DYNEIN BETA CHAIN, CIL 4.58e+02
997 5.0 4485 1 DYHG_CHLRE DYNEIN GAMMA CHAIN, FL 4.58e+02
998 5.0 4563 1 APB_HUMAN APOLIPOPROTEIN B-100 P 4.58e+02
999 5.0 4568 1 DYHC_CAEEL DYNEIN HEAVY CHAIN, CY 4.58e+02
1000 5.0 4660 1 LRP2_RAT LOW-DENSITY LIPOPROTEI 4.58e+02

ALIGNMENTS

RESULT 1
ID CVN NOSEL STANDARD; PRT; 101 AA.
AC P81180;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE CYANOVIRIN-N (CV-N).
OS Nostoc ellipsosporum.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
RN [1]
RP SEQUENCE.
RX MEDLINE; 97445156.
RA Gustafson K.R., Sowder R.C. II, Henderson L.E., Cardellina J.H. II,
RA McMahon J.B., Rajamani U., Pannell L.K., Boyd M.R.;
RA "Isolation, primary sequence determination, and disulfide bond
structure of cyanovirin-N, an anti-HIV (human immunodeficiency virus)
protein from the cyanobacterium Nostoc ellipsosporum.";
RA Biochem. Biophys. Res. Commun. 238:223-228(1997).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
RX MEDLINE; 99262850.
RA Yang F., Bewley C.A., Louis J.M., Gustafson K.R., Boyd M.R.,
RA Gronenborn A.M., Clore G.M., Wlodawer A.;
RA "Crystal structure of cyanovirin-N, a potent HIV-inactivating protein,
shows unexpected domain swapping.";
RA J. Mol. Biol. 288:403-412(1999).
CC -1- FUNCTION: CAPABLE OF INHIBITING HIV-1 AND HIV-2 INFECTION AND
CC REPLICATION.
CC -1- MISCELLANEOUS: CLEAVAGE OF THE DISULFIDE BONDS RESULTS IN THE LOSS
CC OF ANTI-HIV ACTIVITY.
CC PDB; 3EZM; 23-DEC-98.
KW Antiviral; Protein synthesis inhibitor; 3D-structure.
FT DISULFID 8 22
FT DISULFID 58 73
SQ SEQUENCE 101 AA; 11013 MW; 1F84E5B886CCE973 CRC64;

Query Match 59.4%; Score 60; DB 1; Length 101;
Best Local Similarity 100.0%; Pred. No. 5.03e-177;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 42 NVDSGLKQPSNFETCRNTQLAGSSELAACKTRAAQOFVSTKINLDDHIANIDGTLKYE 101

QY 42 NVDSGLKQPSNFETCRNTQLAGSSELAACKTRAAQOFVSTKINLDDHIANIDGTLKYE 101
RESULT 2
ID DORS_DROME STANDARD; PRT; 678 AA.
AC P15330;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE EMBRYONIC POLARITY DORSAL PROTEIN.
DL.
GN Drosophila melanogaster (Fruit fly).
OS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 88042799.
RA Steward R.;
RT "Dorsal, an embryonic polarity gene in Drosophila, is homologous to
RT the vertebrate proto-oncogene, c-rel.";
RL Science 238:692-694(1987).
RN [2]
RP SUBCELLULAR LOCATION, AND REVISIONS.
RX MEDLINE; 90090617.
RA Steward R.;
RT "Relocalization of the dorsal protein from the cytoplasm to the
RT nucleus correlates with its function.";
RL Cell 59:1179-1188(1989).
CC -1- FUNCTION: EMBRYONIC DEVELOPMENTAL PROTEIN. THE LATERAL OR VENTRAL
CC IDENTITY OF A CELL DEPENDS UPON THE CONCENTRATION OF DORSAL
CC PROTEIN IN ITS NUCLEUS DURING THE BLASTODERM STAGE. DORSAL IS A
CC MORPHOGENETIC PROTEIN THAT SPECIFICALLY BINDS TO THE KAPPA B-
CC RELATED CONSENSUS SEQUENCE 5'-GRGAAACC-3', LOCATED IN THE
CC ENHANCER REGION OF ZYGOTIC GENES SUCH AS ZEN, TWIST, SNAIL AND
CC DECAPENTAPLEGIC.
CC -1- SUBCELLULAR LOCATION: IN VENTRAL REGIONS IT IS FIRST CYTOPLASMIC,
CC THEN THE PROTEIN IS RELOCALIZED IN THE NUCLEUS. ITS NUCLEAR
CC LOCALIZATION IS ESSENTIAL TO ITS FUNCTION AS A MORPHOGEN. IN
CC DORSAL REGIONS IT REMAINS CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE REL/DORSAL FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M23702; AAA28479.1; -.
CC PIR; A30350; A30350.
CC HSP; P25799; IBFS.
CC TRANSFAC; T00196; -.
CC FLYBASE; FBgn0000462; dl.
CC PRAM; PF00554; RHD; 1.
CC PRINTS; PR00057; NFKB1NSCPFCT.
CC PROSITE; PS01204; REL_1; 1.
KW Developmental protein; Nuclear protein; Phosphorylation.
FT DOMAIN 47 342 REL-LIKE (RHD).
FT DOMAIN 335 340 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT MOD_RES 312 312 PHOSPHORYLATION (BY CAKP) (POTENTIAL).
SQ SEQUENCE 678 AA; 75475 MW; C23630F44D5FCBAF CRC64;

Query Match 7.9%; Score 8; DB 1; Length 678;
Best Local Similarity 100.0%; Pred. No. 4.79e-04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 181 SSIDLNSV 188

QY 32 SSIDLNSV 39


```
RESULT 3
ID PRC6_LYCES STANDARD; PRT; 259 AA.
AC O24030;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE PROTEASOME ALPHA SUBUNIT (EC 3.4.99.46) (MULTICATALYTIC ENDOPEPTIDASE
GN COMPLEX ALPHA SUBUNIT).
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Asteridae; euasterids I; Solanales;
OC Solanaceae; Solanum.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. LUKULLUS;
RA Ziethe K., Stenzel I., Hertel S.C., Koeck M.;
RT Cloning and characterization of PSR5, a tomato cDNA encoding a 20S
subunit from the proteasome repressed by phosphate starvation.*;
(In) Plant Gene Register PGR98-065.
CC -!- FUNCTION: THE PROTEASOME IS A MULTICATALYTIC PROTEINASE COMPLEX
WHICH IS CHARACTERIZED BY ITS ABILITY TO CLEAVE PEPTIDES WITH
ARG, PHE, TYR, LEU, AND GLU ADJACENT TO THE LEAVING GROUP AT
NEUTRAL OR SLIGHTLY BASIC PH. THE PROTEASOME HAS AN ATP-DEPENDENT
PROTEOLYTIC ACTIVITY.
CC -!- PATHWAY: IS INVOLVED IN AN ATP/UBIQUITIN-DEPENDENT NON-LYSOSOMAL
PROTEOLYTIC PATHWAY.
CC -!- SUBUNIT: THE PROTEASOME IS COMPOSED OF AT LEAST 15 NON IDENTICAL
SUBUNITS WHICH FORM A HIGHLY ORDERED RING-SHAPED STRUCTURE.
CC -!- SUBCELLULAR LOCATION: PROTEASOMES ARE FOUND IN THE CYTOPLASM AND
ALSO IN THE NUCLEUS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY T1A; ALSO KNOWN AS THE
PROTEASOME A-TYPE FAMILY. PROS28 SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; Y14339; CAA74725.1; -
DR HSSP; P25156; LPNA.
DR PFAM; PF00227; proteasome; 1.
DR PROSITE; PS00388; PROTEASOME_A; 1.
K PROTEASOME; Hydrolase; Protease.
P SEQUENCE: 259 AA; 28480 MW; 700B3638C9F0FB52 CRC64;
Query Match 6.9%; Score 7; DB 1; Length 259;
Best Local Similarity 100.0%; Pred. No. 8.39e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 65 NLDDHIA 71
QY 86 NLDDHIA 92
RESULT 4
ID MODA_MYCTU STANDARD; PRT; 261 AA.
AC P95157; O05125;
DT 15-FEB-2000 (Rel. 39, Created)
DT 15-FEB-2000 (Rel. 39, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE MOLYBDATE-BINDING PROTEIN PRECURSOR.
GN MODA OR RV1857 OR MTCY359.16C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RA Asturias J.A., Eltis L.D., Prucha M., Timmis K.N.;
RT
```

```
RA Laqueyrie A.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RA MEDLINE; 98295987.
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
RA Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence.*;
RL Nature 393:537-544(1998).
CC -!- FUNCTION: INVOLVED IN THE TRANSPORT OF MOLYBDENUM INTO THE CELL.
PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM MODABCD.
-!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID
ANCHOR (PROBABLE).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; X99258; CAA67642.1; -
DR EMBL; Z83859; CAB06130.1; -
DR TUBERCULIST; RV1857; -
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Transport; Molybdenum; Membrane; Lipoprotein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 261
FT LIPID 22 22
FT CONFLICT 1 59
FT CONFLICT 100 134
FT CONFLICT 100 134
FT IGCQGGGVAGRSSDKLRHQHDGHRRCRRQSQEDP (IN
REF. 1).
SQ SEQUENCE 261 AA; 26576 MW; CFE292F6D595A9F5 CRC64;
Query Match 6.9%; Score 7; DB 1; Length 261;
Best Local Similarity 100.0%; Pred. No. 8.39e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 74 AGSSELA 80
QY 64 AGSSELA 70
RESULT 5
ID BHCL_RHOQ STANDARD; PRT; 291 AA.
AC P47231;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE BIPHENYL-2,3-DIOL 1,2-DIOXYGENASE I (EC 1.13.11.39) (23OHBP
DE OXYGENASE I) (2,3-DIHYDROXYBIPHENYL DIOXYGENASE I) (DHBD I).
GN BPHCL.
OS Rhodococcus globulus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Actinobacteriaceae; Rhodococcu.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=P6;
RA MEDLINE; 94171820.
RA Asturias J.A., Eltis L.D., Prucha M., Timmis K.N.;
```

RT Analysis of three 2,3-dihydroxybiphenyl 1,2-dioxygenases found in
 RT Rhodococcus globerulus P6. Identification of a new family of
 RL extradiol dioxygenases.";
 J. Biol. Chem. 269:7807-7815(1994).
 CC -1- CATALYTIC ACTIVITY: BIPHENYL-2,3-DIOL + O(2) = 2-HYDROXY-6-OXO-
 CC 6-PHENYLHEXA-2,4-DIENOATE + H(2O).
 CC -1- COFACTOR: FERROUS ION.
 CC -1- PATHWAY: DEGRADATION OF BIPHENYLS AND POLYCHLOROBIPHENYLS (PCB) TO
 CC BENZOIC ACID AND CHLOROBENZOIC ACIDS.
 CC -1- SIMILARITY: BELONGS TO THE EXTRADIOL RING-CLEAVAGE DIOXYGENASE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X75633; CAA53297.1; -;
 CC HSP; P47228; LHAN.
 CC PFAM; PF01013; Extradiol_dioxy; 1.
 DR PROSITE; PS00082; EXTRADIOL_DIOXYGENAS; 1.
 KW Oxidoreductase; Dioxygenase; Aromatic hydrocarbons catabolism; Iron.
 FT METAL 146 146
 FT METAL 210 210 IRON (BY SIMILARITY).
 FT METAL 260 260 IRON (BY SIMILARITY).
 SQ SEQUENCE 291 AA; 32081 MW; 104F189FE1EDDAA CRC64;
 Query Match 6.9%; Score 7; DB 1; Length 291;
 Best Local Similarity 100.0%; Pred. No. 8.39e-02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 91 SSELAAE 97
 QY 66 SSELAAE 72
 RESULT 6
 ID OSTG_YEAST STANDARD; PRT; 350 AA.
 AC P48439;
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE DOLICHYL-DIPHOSPHOOLIGOSACCHARIDE--PROTEIN GLYCOSYLTRANSFERASE GAMMA
 DE SUBUNIT PRECURSOR (EC 2.4.1.119) (OLIGOSACCHARYL TRANSFERASE GAMMA
 DE SUBUNIT) (OLIGOSACCHARYL TRANSFERASE 34 KDA SUBUNIT).
 CC OST3 OR YOR085W OR YOR3124W.
 CC Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
 CC Saccharomycetaceae; Saccharomycetes.
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-37; 99-103; 123-153 & 345-350.
 RX MEDLINE; 95348180.
 RA Karaoglu D., Kelleher D.J., Gilmore R.;
 RT "Functional characterization of Ost3p. Loss of the 34-kD subunit of
 RT the Saccharomyces cerevisiae oligosaccharyltransferase results in
 RT biased underglycosylation of acceptor substrates.";
 RN J. Cell Biol. 130:567-577(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 97344368.
 RA Voss H., Benes V., Andrade M.A., Valencia A., Rechmann S., Teodoru C.,
 RA Schwager C., Paces V., Sander C., Ansorge W.;
 RT "DNA sequencing and analysis of 130 kb from yeast chromosome XV.";
 RN Yeast 13:653-672(1997).
 CC -1- FUNCTION: MAY ENHANCE OLIGOSACCHARIDE TRANSFER IN VIVO TO A SUBSET
 CC OF ACCEPTOR SUBSTRATES.
 CC -1- CATALYTIC ACTIVITY: DOLICHYL DIPHOSPHOOLIGOSACCHARIDE + PROTEIN
 CC L-ASPARAGINE = DOLICHYL DIPHOSPHATE + A GLYCOPROTEIN WITH THE
 CC OLIGOSACCHARIDE CHAIN ATTACHED BY GLYCOSYLAMINE LINKAGE TO PROTEIN
 CC L-ASPARAGINE.

CC -1- PATHWAY: GLYCOSYLATION.
 CC -1- SUBUNIT: YEAST OST SEEMS TO CONSIST OF SIX DIFFERENT SUBUNITS
 CC (ALPHA TO ZETA).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. ENDOPLASMIC
 CC RETICULUM (PROBABLE).
 CC -1- SIMILARITY: BELONGS TO THE OST3 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U25052; AAC49042.1; -;
 CC EMBL; X94335; CAA64007.1; -;
 CC EMBL; Z74993; CAA99280.1; -;
 CC SGD; L0002943; OST3.
 KW Transferase; Endoplasmic reticulum; Transmembrane; Signal.
 FT SIGNAL 1 22
 FT CHAIN 23 350 DOLICHYL-DIPHOSPHOOLIGOSACCHARIDE--
 FT SUBUNIT. PROTEIN GLYCOSYLTRANSFERASE GAMMA
 FT TRANSMEM 186 203 POTENTIAL.
 FT TRANSMEM 218 237 POTENTIAL.
 FT TRANSMEM 272 288 POTENTIAL.
 FT TRANSMEM 310 329 POTENTIAL.
 SQ SEQUENCE 350 AA; 39483 MW; EEA573BC14E99380 CRC64;
 Query Match 6.9%; Score 7; DB 1; Length 350;
 Best Local Similarity 100.0%; Pred. No. 8.39e-02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 241 RNTOLAG 247
 QY 59 RNTOLAG 65
 RESULT 7
 ID PUR8_HELPJ STANDARD; PRT; 440 AA.
 AC Q9ZKA2;
 DT 15-FEB-2000 (Rel. 39, Created)
 DT 15-FEB-2000 (Rel. 39, Last sequence update)
 DT 15-FEB-2000 (Rel. 39, Last annotation update)
 DE ADENYLOSUCCLINATE LYASE (EC 4.3.2.2) (ADENYLOSUCCLINASE) (ASL).
 GN PURB OR HP1112.
 OS Helicobacter pylori J99 (Campylobacter pylori J99).
 CC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 CC Helicobacter.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-J99;
 RX MEDLINE; 99120557.
 RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
 RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
 RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
 RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
 RA Trust T.J.;
 RT "Genomic sequence comparison of two unrelated isolates of the human
 RT gastric pathogen Helicobacter pylori.";
 RN Nature 397:176-180(1999).
 CC -1- CATALYTIC ACTIVITY: 1-(5-PHOSPHORIBOSYL)-4-(N-SUCCINO-CARBOXAMIDE)
 CC -5-AMINOIMIDAZOLE = FUMARATE + 5'-PHOSPHORIBOSYL-5-AMINO-4-
 CC IMIDAZOLECARBOXAMIDE (ALSO CATALYZES: N6-(1,2-DICARBOXYETHYL)AMP =
 CC FUMARATE + AMP).
 CC -1- PATHWAY: EIGHT STEP IN DE NOVO PURINE BIOSYNTHESIS.
 CC -1- SIMILARITY: BELONGS TO THE LYASE 1 FAMILY. ADENYLOSUCCLINATE LYASE
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: AF001531; AAD06609.1; -
 DR PFAM: PF00206; lyase_1; 1.
 DR PROSITE: PS00163; FUMARATE_LYASES; 1.
 KW Purine biosynthesis; Lyase.
 SQ SEQUENCE 440 AA; 49867 MW; C67D3C02AEFA2EFE CRC64;

Query Match 6.9%; Score 7; DB 1; Length 440;
 Best Local Similarity 100.0%; Pred. No. 8.39e-02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 327 LNSVIEV 333

QY 36 LNSVIEV 42

|||||

LT 8
 ID DNAK_ALCEU STANDARD; PRT: 656 AA.
 AC O33522;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DE DNAK PROTEIN (HEAT SHOCK PROTEIN 70) (HSP70).
 GN DNAK.
 OS Alkaligenes eutrophus (Ralstonia eutropha).
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
 OC Ralstonia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CH34;
 RA Talbi S., van der Lelie D.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: ACTS AS A CHAPERONE (BY SIMILARITY).
 CC -!- INDUCTION: BY STRESS CONDITIONS E.G. HEAT SHOCK (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: AJ001727; CAA04955.1; -
 DR HSP: P04775; 1DKX.
 DR PFAM: PF00012; HSP70; 1.
 DR PRINTS: PR00301; HEATSHOCK70.
 DR PROSITE: PS00297; HSP70.1; 1.
 DR PROSITE: PS00329; HSP70.2; 1.
 DR PROSITE: PS01036; HSP70.3; 1.
 KW Chaperone; ATP-binding; Heat shock.
 SQ SEQUENCE 656 AA; 71300 MW; 5C7D5D2CE22F5F97 CRC64;

Query Match 6.9%; Score 7; DB 1; Length 656;
 Best Local Similarity 100.0%; Pred. No. 8.39e-02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 377 AIOGSVL 383

QY 12 AIOGSVL 18

|||||

RESULT 9
 ID YDBH_ECOLI STANDARD; PRT: 879 AA.
 AC P52645; P77502; P76855;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-FEB-2000 (Rel. 39, Last annotation update)

DE HYPOTHETICAL 96.8 KDA PROTEIN IN LDHA-TYNA INTERGENIC REGION.
 GN YDBH.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE: 97426617.
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner F.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE: 97251357.

RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
 RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
 RA Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,
 RA Samede G., Seki Y., Sivasubram S., Tagami H., Takeda J.,
 RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
 RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
 corresponding to the 28.0-40.1 min region on the linkage map.";
 RL DNA Res. 3:363-377(1996).
 RN [3]
 RP SEQUENCE OF 1-331 FROM N.A.
 RC STRAIN=K12;
 RA Bunch P.K., Mat-Jan F., Lee N.A., Deayala B.A., Clark D.P.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.

RN [4]
 RP IDENTIFICATION.
 RA Rudd K.E.;
 RL Unpublished observations (MAR-1996).
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: AE000235; AAC74463.1; -
 DR EMBL: D90776; CAB20902.1; -
 DR EMBL: D90777; CAB20910.1; -
 DR EMBL: U36928; -; NOT_ANNOTATED_CDS.
 DR ECOGENE: EG13180; ydbH.
 KW Hypothetical protein.
 FT CONFLICT 36 36 I -> L (IN REF. 3).
 SQ SEQUENCE 879 AA; 96834 MW; 43892C8391751C1D CRC64;

Query Match 6.9%; Score 7; DB 1; Length 879;
 Best Local Similarity 100.0%; Pred. No. 8.39e-02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 801 STKINLD 807

QY 82 STKINLD 88

|||||

RESULT 10
 ID PER3_MOUSE STANDARD; PRT: 1113 AA.
 AC O70361;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE PERIOD CIRCADIAN PROTEIN 3 (MPER3).
 GN PER3.
 OS Mus musculus (Mouse).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=BRAIN;
 RX MEDLINE; 98318231.
 RA Zylka M.J., Shearman L.P., Weaver D.R., Reppert S.M.;
 RT "Three period homologs in mammals: differential light responses in the
 RT suprachiasmatic circadian clock and oscillatin transcripts outside of
 RT brain."
 RL Neuron 20:1103-1110(1998).
 CC -|- FUNCTION: CIRCADIAN REGULATOR THAT MAY ACT AS A TRANSCRIPTION
 CC FACTOR. BEHAVES AS A NEGATIVE ELEMENT IN CIRCADIAN TRANSCRIPTIONAL
 CC LOOP. DOES NOT APPEAR TO BIND DNA, SUGGESTING INDIRECT
 CC TRANSCRIPTIONAL INHIBITION (BY SIMILARITY). EXPRESSION OSCILLATES
 CC IN THE SUPRACHIASMATIC NUCLEI (SCN) AND EYES. THE EXPRESSION
 CC RHYTHMS APPEAR TO ORIGINATE FROM RETINA.
 CC -|- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
 CC -|- TISSUE SPECIFICITY: WIDELY EXPRESSED. EXPRESSED IN HEART, BRAIN,
 CC LUNG, LIVER, SKELETAL MUSCLE, TESTIS, AND AT LOW LEVEL IN THE
 CC SPLEEN AND KIDNEY. IN BRAIN, MAINLY FOUND IN THE SCN. HIPPOCAMPUS,
 CC PIRIFORM CORTEX, AND CEREBELLUM. LOWER LEVEL OF EXPRESSION IN THE
 CC NEOCORTEX. EXPRESSION EXHIBITS SYNCHRONOUS OSCILLATIONS IN LIVER,
 CC SKELETAL MUSCLE AND TESTIS.
 CC -|- INDUCTION: NOT ACUTELY INDUCED BY LIGHT IN SCN DURING SUBJECTIVE
 CC NIGHT.
 CC -|- SIMILARITY: CONTAINS A PAS (PER-ARNT-SIM) DIMERIZATION DOMAIN.
 CC -|- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 CC TRANSCRIPTION FACTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF050182; AAC40147.1; -
 DR GMD; MGI:1277134; PER3.
 DR PFAM; PF00989; PAS; 1.
 KW Transcription regulation; Nuclear protein; Repeat; Biological rhythms.
 FT DOMAIN 52 86
 FT REPEAT 121 187 PAS-1.
 FT REPEAT 259 323 PAS-2.
 FT DOMAIN 336 376 PAC MOTIF.
 FT DOMAIN 562 565 POLY-SER.
 SQ SEQUENCE 1113 AA; 120939 MW; 8121E235D100A627 CRC64;
 Query Match 6.9%; Score 7; DB 1; Length 1113;
 Best Local Similarity 100.0%; Pred. No. 8.39e-02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 863 AGSSELA 869
 QY 64 AGSSELA 70
 RESULT 11
 ID ICDB_PSEX STANDARD; PRT; 1148 AA.
 AC O30611;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE ICE NUCLEATION PROTEIN.
 GN INAK.
 OS Pseudomonas syringae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonas group;
 CC Pseudomonas.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KCTC 1832;
 RA Jung H.-C., Pan J.-G.;

RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 CC -|- FUNCTION: ICE NUCLEATION PROTEINS ENABLE BACTERIA TO NUCLEATE
 CC CRYSTALLIZATION IN SUPERCOOLED WATER.
 CC -|- SUBCELLULAR LOCATION: OUTER MEMBRANE (BY SIMILARITY).
 CC -|- DOMAIN: CONTAINS MANY IMPERFECT REPEATS OF THE CONSENSUS
 CC OCTAPEPTIDE A-G-Y-G-S-T-L-T; FURTHERON A 16-RESIDUE AND A
 CC REGIONAL 48-RESIDUE PERIODICITY IS SUPERIMPOSED.
 CC -|- SIMILARITY: BELONGS TO THE BACTERIAL ICE NUCLEATION PROTEIN
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF013159; AAB66891.1; -
 DR PFAM; PF00818; Ice_nucleation; 57.
 DR PRINTS; PR00327; ICENUCLEATN
 DR PROSITE; PS00314; ICE_NUCLEATION; 38.
 KW Ice nucleation; Repeat; Outer membrane.
 FT DOMAIN 161 164 POLY-ALA.
 FT DOMAIN 180 1099 OCTAPEPTIDE PERIODICITY.
 SQ SEQUENCE 1148 AA; 113811 MW; BA4019CF20FAE224 CRC64;
 Query Match 6.9%; Score 7; DB 1; Length 1148;
 Best Local Similarity 100.0%; Pred. No. 8.39e-02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 534 QGSVLTS 540
 QY 14 QGSVLTS 20
 RESULT 12
 ID ICBV_PSEX STANDARD; PRT; 1196 AA.
 AC O33479;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE ICE NUCLEATION PROTEIN.
 GN INAV.
 OS Pseudomonas syringae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonas group;
 CC Pseudomonas.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=INA5;
 RX MEDLINE; 97462815.
 RA Schmid D., Fridmore D., Capitani G., Battistutta R., Neeser J.-R.,
 RA Jann A.;
 RT "Molecular organisation of the ice nucleation protein Inav from
 RT Pseudomonas syringae."
 RT FEBS Lett. 414:590-594(1997).
 CC -|- FUNCTION: ICE NUCLEATION PROTEINS ENABLE BACTERIA TO NUCLEATE
 CC CRYSTALLIZATION IN SUPERCOOLED WATER.
 CC -|- SUBUNIT: MEMBRANE ENVIRONMENT OR AGGREGATION SEEMS TO BE REQUIRED
 CC FOR ICE NUCLEATION ACTIVITY.
 CC -|- SUBCELLULAR LOCATION: OUTER MEMBRANE (BY SIMILARITY).
 CC -|- DOMAIN: CONTAINS MANY IMPERFECT REPEATS OF THE CONSENSUS
 CC OCTAPEPTIDE A-G-Y-G-S-T-L-T; FURTHERON A 16-RESIDUE AND A
 CC REGIONAL 48-RESIDUE PERIODICITY IS SUPERIMPOSED.
 CC -|- SIMILARITY: BELONGS TO THE BACTERIAL ICE NUCLEATION PROTEIN
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

FT CONFLICT 229 229 E -> EE (IN REF. 2).
 FT CONFLICT 535 535 V -> M (IN REF. 2).
 FT CONFLICT 546 547 WN -> RT (IN REF. 2).
 FT CONFLICT 574 574 MISSING (IN REF. 2).
 FT CONFLICT 591 591 T -> A (IN REF. 2).
 FT CONFLICT 621 621 T -> N (IN REF. 2).
 FT CONFLICT 1295 1295 V -> L (IN REF. 2).
 SQ SEQUENCE 1416 AA; 158365 MW; 447C8110A775DD42 CRC64;

Query Match 6.9%; Score 7; DB 1; Length 1416;
 Best Local Similarity 100.0%; Pred. No. 8.39e-02; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0;

.Db 527 GSVLTST 533
 |||||
 QY 15 GSVLTST 21

RESULT 15
 ID Y076_NPVAC STANDARD; PRT; 84 AA.
 Q06690;
 01-NOV-1995 (Rel. 32, Created)
 01-NOV-1995 (Rel. 32, Last sequence update)
 01-NOV-1997 (Rel. 35, Last annotation update)
 DE HYPOTHETICAL 9.4 KDA PROTEIN IN IAP2-VLF1 INTERGENIC REGION.
 OS Autographa californica nuclear polyhedrosis virus (ACMNPV).
 CC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 CC Nucleopolyhedrovirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C6;
 RX MEDLINE; 94303173.
 RA Ayres M.D., Howard S.C., Kuzio J., Lopez-Perber M., Possee R.D.;
 RA "The complete DNA sequence of Autographa californica nuclear
 RT polyhedrosis virus";
 RL Virology 202:586-605(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=E2;
 RX MEDLINE; 94172322.
 RA Kool M., Broer R., Zuidema D., Goldbach R.W., Vlak J.M.;
 RA "Nucleotide sequence and genetic organization of a 7.3 kb region (map
 RT unit 47 to 52.5) of Autographa californica nuclear polyhedrosis virus
 RT fragment EcoRI-C";
 RL J. Gen. Virol. 75:487-494(1994).
 CC -1- SIMILARITY: TO CORRESPONDING ORF IN OPMNPV.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; L22858; AAA66706.1; -;
 DR EMBL; X71415; CAA50539.1; -;
 DR PIR; S36691; S36691
 KW Hypothetical protein.
 FT CONFLICT 61 61 S -> N (IN REF. 2).
 FT CONFLICT 71 71 A -> T (IN REF. 2).
 FT CONFLICT 74 74 A -> S (IN REF. 2).
 SQ SEQUENCE 84 AA; 9440 MW; AB9E33C41C914439 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 84;
 Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

.Db 79 NTSSID 84
 |||||
 QY 30 NTSSID 35

RESULT 16
 ID RL44_YEAST STANDARD; PRT; 105 AA.
 AC P02405;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE 60S RIBOSOMAL PROTEIN L42 (L44) (YL27) (YP44) (L41).
 GN (RPL42A OR RPL44 OR SCL41A OR RPL41A OR YNL162W OR NL722) AND
 GN (RPL42B OR RPL44 OR SCL41B OR RPL41B OR MAK18 OR YHR141C).
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
 OC Saccharomycetaceae; Saccharomycetes.
 RN [1]
 RP SEQUENCE FROM N.A. (RPL42A AND RPL42B).
 RX MEDLINE; 92104971.
 RA Kawai S., Murao S., Mochizuki M., Shibuya I., Yano K., Takagi M.;
 RA "Drastic alteration of cycloheximide sensitivity by substitution of
 RT one amino acid in the L41 ribosomal protein of yeasts";
 RL J. Bacteriol. 174:254-262(1992).
 RN [2]
 RP SEQUENCE FROM N.A. (RPL42A).
 RC STRAIN=S288C / FY1679;
 RX MEDLINE; 96287653.
 RA Nasr F., Becam A.-M., Herbert C.J.;
 RA "The sequence of 36.8 kb from the left arm of chromosome XIV reveals
 RT 24 complete open reading frames: 18 correspond to new genes, one of
 RT which encodes a protein similar to the human myotonic dystrophy
 RT kinase";
 RL Yeast 12:169-175(1996).
 RN [3]
 RP SEQUENCE FROM N.A. (RPL42B).
 RC STRAIN=S288C / AB972;
 RX MEDLINE; 94378003.
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
 RA Du Z., Favellio A., Fulton L., Gattung S., Geisel C., Kirsten J.,
 RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
 RA Latreille P., Louis E.J., Macri C., Mardis E., Meneses S., Mouser L.,
 RA Nhan M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,
 RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
 RA Vaudin M.;
 RA "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
 RT VIII";
 RL Science 265:2077-2082(1994).
 RN [4]
 RP SEQUENCE.
 RX MEDLINE; 79086263.
 RA Itoh T., Wittmann-Liebold B.;
 RA "The primary structure of protein 44 from the large subunit of yeast
 RL ribosomes";
 RN FEBS Lett. 96:399-402(1978).
 RN [5]
 RP IDENTIFICATION AS MAK18.
 RX MEDLINE; 95270609.
 RA Carroll K., Wickner R.B.;
 RA "Translation and M1 double-stranded RNA propagation: MAK18 - RPL41B
 RT and cycloheximide curing";
 RL J. Bacteriol. 177:2887-2891(1995).
 CC -1- PPM: THE RESIDUES AT POSITIONS 40 AND 54 WERE NOT POSITIVELY
 CC IDENTIFIED BUT ARE CHEMICALLY RELATED TO MONOMETHYLLYSINE.
 CC -1- MISCELLANEOUS: THERE ARE TWO GENES FOR L42 IN YEAST.
 CC -1- SIMILARITY: BELONGS TO THE L44E FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D10578; BAA01435.1; -;
 DR EMBL; D10579; BAA01436.1; -;
 DR EMBL; X92517; CAA63277.1; -;

EMBL; 271438; CAA96049.1; ALT_SEQ.

DR EMBL; M62391; AAA34927.1; -

DR EMBL; M62392; AAA34928.1; -

DR EMBL; U10398; AAB68420.1; -

DR PIR; A02781; R6BY44

DR PIR; B43301; B43301

DR PIR; C43301; C43301

DR PIR; S48985; S48985

DR SGD; L0001731; RPL42A

DR SGD; L0000990; RPL42B

DR PFAM; PF00935; Ribosomal_L44; 1

DR PROSITE; PS01172; RIBOSOMAL_L44E; 1

KW Ribosomal protein; Cycloheximide resistance; Multigene family.

FT INIT_MET 0 0 P -> Q (CONFERS RESISTANCE TO

FT VARIANT 55 55 CYCLOHEXIMIDE, AN INHIBITOR OF

FT POLYPEPTIDE ELONGATION).

FT MISSING (IN REF. 4).

FT CONFLICT 87 88 MISSING (IN REF. 4).

FT SEQUENCE 105 AA; 12080 MW; 916634ECE3CB0121 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 105;

Best Local Similarity 100.0%; Pred. No. 8.75e+00;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 76 CKTRAQ 81

QY 73 CKTRAQ 78

RESULT 17

ID HER1_CAEEL STANDARD; PRT; 175 AA.

AC P34704;

DT 01-FEB-1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 15-FEB-2000 (Rel. 39, Last annotation update)

DE HER-1 PROTEIN PRECURSOR.

GN HER-1 OR ZK287.8.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RX MEDLINE; 93170661.

RA Perry M.D., Li W., Trent C., Robertson B., Fire A., Hageman J.M.,

RA Wood W.B.;

RT "Molecular characterization of the her-1 gene suggests a direct role

in cell signaling during Caenorhabditis elegans sex determination.";

Genes Dev. 7:216-228(1993).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RA McMurray A.;

RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: DICTATES MALE DEVELOPMENT. PROBABLY PLAYS A DIRECT ROLE

IN CELL SIGNALING DURING C.ELEGANS SEX DETERMINATION.

CC -!- SUBCELLULAR LOCATION: SECRETED.

CC -!- ALTERNATIVE PRODUCTS: TWO TYPES OF TRANSCRIPTS HAVE BEEN

IDENTIFIED: A LARGER; BIOLOGICALLY ACTIVE PRODUCT AND A SMALLER

TRANSCRIPT THAT PRODUCES A NON ACTIVE PROTEIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to license@isb-sib.ch.

CC -----

DR EMBL; Z19595; CAA79650.1; -

DR EMBL; Z19595; CAA79651.1; -

DR EMBL; Z70757; CAA94804.1; -

DR PIR; A46388; A46388

DR PIR; S32245; S32245.

DR PIR; S32246; S32246.

DR WORMPEP; ZK287.8; CE06617.

KW Developmental protein; Glycoprotein; Signal; Alternative splicing.

FT SIGNAL 1 18 POTENTIAL.

FT CHAIN 19 175 HER-1 PROTEIN.

FT CARBOHYD 98 98 POTENTIAL.

FT CARBOHYD 163 163 POTENTIAL.

FT VARSPLIC 1 111 MISSING (IN TRUNCATED ISOFORM).

SQ SEQUENCE 175 AA; 20172 MW; C330DFE9BB3D869A CRC64;

Query Match 5.9%; Score 6; DB 1; Length 175;

Best Local Similarity 100.0%; Pred. No. 8.75e+00;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 81 KINLDD 86

QY 84 KINLDD 89

RESULT 18

ID YH22_VACCV STANDARD; PRT; 184 AA.

AC P17366;

DT 01-AUG-1990 (Rel. 15, Created)

DT 01-AUG-1990 (Rel. 15, Last sequence update)

DT 01-AUG-1990 (Rel. 15, Last annotation update)

DE HYPOTHETICAL 21.7 KDA HINDIII-C PROTEIN.

OS Vaccinia virus (strain WR)

OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;

OC Orthopoxvirus.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 89073756.

RA Kotwal G.J., Moss B.;

RT "Analysis of a large cluster of nonessential genes deleted from a

vaccinia virus terminal transposition mutant.";

RL Virology 167:524-537(1988).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to license@isb-sib.ch.

CC -----

DR EMBL; M22812; AAA69592.1; -

DR PIR; A31829; WZVZAL.

KW Hypothetical protein; Early protein.

SQ SEQUENCE 184 AA; 21604 MW; E8CCGCCDB5529B293 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 184;

Best Local Similarity 100.0%; Pred. No. 8.75e+00;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 9 NTSSID 14

QY 30 NTSSID 35

RESULT 19

ID HTGA_ECOLI STANDARD; PRT; 196 AA.

AC P28697;

DT 01-DEC-1992 (Rel. 24, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE HEAT SHOCK PROTEIN HTGA (HEAT SHOCK PROTEIN HTPY).

GN HTGA OR HTPY.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.

RN [1]

RP SEQUENCE FROM N.A.

```

RX MEDLINE; 94003405.
RA James R., Dean D.O., Debbage J.;
RT "Five open reading frames upstream of the dnaK gene of E. coli.";
RL DNA Seq. 3:327-332(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / W3110;
RX MEDLINE; 93239687.
RA Missiakas D., Georgopoulos C., Raina S.;
RT "The Escherichia coli heat shock gene htpy: mutational analysis,
cloning, sequencing, and transcriptional regulation.";
RL J. Bacteriol. 175:2613-2624(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE; 92334977.
RA Yura H., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
Isono K., Mirobuchi K., Nakata A.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of
the 0-2.4 min region.";
RL Nucleic Acids Res. 20:3305-3308(1992).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE; 97426617.
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner F.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
CC -1- FUNCTION: NOT KNOWN, REQUIRED FOR HIGH-TEMPERATURE GROWTH OF
E. COLI. IT IS POSSIBLE THAT HTGA PROTEIN SOMEHOW REGULATES EITHER
TRANSCRIPTION OF THE RPOH GENE OR THE ACTIVITY OF ITS GENE PRODUCT
SIGMA-32. PLAYS A ROLE OPPOSITE THAT OF DNAJ, DNAJ, AND GRPE IN
TERMS OF HEAT SHOCK REGULATION, ANTAGONIZING THE NEGATIVE OF
THESE AND THUS FINE-TUNING THE HEAT SHOCK RESPONSE.
CC -1- INDUCTION: INDUCED AT HIGH TEMPERATURES.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-36 IS THE INITIATOR.
CC -1- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
FRAMESHIFT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X67700; CAA47932.1; ALT INIT.
DR EMBL; L03720; AAA23992.1; -.
DR EMBL; L03720; AAA23993.1; ALT INIT.
DR EMBL; D10483; -. NOT_ANNOTATED_CDS.
DR EMBL; AE000112; AAC73123.1; -.
DR PIR; S28460; S28460.
DR ECGENE; EG11509; HTGA.
KW Heat shock.
FT DOMAIN 57 65 POLY-SER.
SQ SEQUENCE 196 AA; 21225 MW; ECA6154160A40993 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 196;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 159 SIDLNS 164
Qy 33 SIDLNS 38
|||||

RESULT 20 STANDARD; PRT; 198 AA.
ID VIF_BIV27
VIF_VIF27 5.9%; Score 6; DB 1; Length 198;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

AC P19562;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE VIRION INFECTIVITY FACTOR (Q PROTEIN).
GN VIF.
OS Bovine immunodeficiency virus (isolate 127) (BIV).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90223985.
RA Garvey K.J., Oberste M.S., Elser J.E., Braun M.J., Gonda M.A.;
RT "Nucleotide sequence and genome organization of biologically active
proviruses of the bovine immunodeficiency-like virus.";
RL Virology 175:391-409(1990).
CC -1- FUNCTION: DETERMINES VIRUS INFECTIVITY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M32691; -. NOT_ANNOTATED_CDS.
DR PIR; M32691; QORF$BIV106.
SQ SEQUENCE 198 AA; 22827 MW; 119F49C5FB898529 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 198;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ID VIF_BIV06 STANDARD; PRT; 198 AA.
AC P19562;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE VIRION INFECTIVITY FACTOR (Q PROTEIN).
GN VIF.
OS Bovine immunodeficiency virus (isolate 106) (BIV).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90223985.
RA Garvey K.J., Oberste M.S., Elser J.E., Braun M.J., Gonda M.A.;
RT "Nucleotide sequence and genome organization of biologically active
proviruses of the bovine immunodeficiency-like virus.";
RL Virology 175:391-409(1990).
CC -1- FUNCTION: DETERMINES VIRUS INFECTIVITY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M32691; -. NOT_ANNOTATED_CDS.
DR PIR; M32691; QORF$BIV106.
SQ SEQUENCE 198 AA; 22827 MW; 119F49C5FB898529 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 198;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 141 SVLTST 146
Qy 16 SVLTST 21
|||||

Query Match 5.9%; Score 6; DB 1; Length 198;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ID VIF_BIV06 STANDARD; PRT; 198 AA.
AC P19562;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE VIRION INFECTIVITY FACTOR (Q PROTEIN).
GN VIF.
OS Bovine immunodeficiency virus (isolate 106) (BIV).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90223985.
RA Garvey K.J., Oberste M.S., Elser J.E., Braun M.J., Gonda M.A.;
RT "Nucleotide sequence and genome organization of biologically active
proviruses of the bovine immunodeficiency-like virus.";
RL Virology 175:391-409(1990).
CC -1- FUNCTION: DETERMINES VIRUS INFECTIVITY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M32691; -. NOT_ANNOTATED_CDS.
DR PIR; M32691; QORF$BIV106.
SQ SEQUENCE 198 AA; 22827 MW; 119F49C5FB898529 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 198;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
Db 141 SVLTST 146
    |||||
Oy 16 SVLTST 21

RESULT 22
ID GCH1_HAEIN STANDARD; PRT; 218 AA.
AC P43866;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE GTP CYCLOHYDROLASE I (EC 3.5.4.16) (GTP-CH-I).
GN FOLE OR H11447.

OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=RD / KW20;
RX MEDLINE; 95350630.
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae RA."
RL Science 269:496-512(1995).
RC -1- CATALYTIC ACTIVITY: GTP + 2 H(2)O -> FORMATE + 2-AMINO-4-HYDROXY-
CC 6-(ERYTHRO-1,2,3-TRIHIDROXYPROPYL)DIHYDROPTERIDINE TRIPHOSPHATE.
CC -1- PATHWAY: FIRST STEP IN THE BIOSYNTHESIS OF TETRAHYDROFOLATE.
CC -1- SUBUNIT: HOMOPOLYMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GTP CYCLOHYDROLASE I FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; U32823; AAC23097.1; -.
LR TIGR; H11447; -.
DR PFAM; PF01227; GTP_cyclohydrol; 1.
DR PROSITE; PS00859; GTP_CYCLOHYDROL_1_1; 1.
DR PROSITE; PS00860; GTP_CYCLOHYDROL_1_2; 1.
KW One-carbon metabolism; Hydrolase; Allosteric enzyme.
FT DISULFID 109 180 BY SIMILARITY.
SQ SEQUENCE 218 AA; 24945 MW; BCC60F0038158D38 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 218;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 105 LTSTCE 110
    |||||
Oy 18 LTSTCE 23

RESULT 23
ID PAA4_ECOLI STANDARD; PRT; 219 AA.
AC P22996;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE RESOLVASE.
GN PARA.

OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=K12;
RX MEDLINE; 92172284.
RA Katzenmeier G., Schmid C., Kellermann J., Lottspeich F., Bacher A.;
RT "Biosynthesis of tetrahydrofolate. Sequence of GTP cyclohydrolase I
RT from Escherichia coli."
RL Biol. Chem. Hoppe-Seyler 372:991-997(1991).
RN [2]
RN REVISIONS.
RX MEDLINE; 94136208.
RA Schmid C., Meining W., Weinkauff S., Bachmann L., Ritz H.,
RA Eberhardt S., Gimbel W., Werner T., Lahm H.W., Nar H., Bacher A.;
RT "Studies on GTP cyclohydrolase I of Escherichia coli."
RL Adv. Exp. Med. Biol. 338:157-162(1993).
RN [3]
RN SEQUENCE FROM N.A.

OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX STRAIN=K12;
RX MEDLINE; 92172284.
RA Katzenmeier G., Schmid C., Kellermann J., Lottspeich F., Bacher A.;
RT "Biosynthesis of tetrahydrofolate. Sequence of GTP cyclohydrolase I
RT from Escherichia coli."
RL Biol. Chem. Hoppe-Seyler 372:991-997(1991).
RN [2]
RN REVISIONS.
RX MEDLINE; 94136208.
RA Schmid C., Meining W., Weinkauff S., Bachmann L., Ritz H.,
RA Eberhardt S., Gimbel W., Werner T., Lahm H.W., Nar H., Bacher A.;
RT "Studies on GTP cyclohydrolase I of Escherichia coli."
RL Adv. Exp. Med. Biol. 338:157-162(1993).
RN [3]
RN SEQUENCE FROM N.A.
```


FT SITE 100 INVOLVED IN RECEPTOR BINDING.
FT SITE 103 INVOLVED IN RECEPTOR BINDING.
FT DISULFID 89 BY SIMILARITY.
FT DISULFID 122 BY SIMILARITY.
FT DISULFID 126 BY SIMILARITY.
FT DISULFID 116 INTERCHAIN (BY SIMILARITY).
FT DISULFID 125 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 55 POTENTIAL.
FT NON_TER 225
SQ SEQUENCE 225 AA; 25603 MW; 0DAE138B0AA70F0F CRC64;

Query Match 5.9%; Score 6; DB 1; Length 225;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 87 ABEKTR 92
|||||
QY 71 ABEKTR 76

BLT 26
ID TSIS SMSAV STANDARD; PRT; 226 AA.
AC P01128; O41283;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PDGF-RELATED TRANSFORMING PROTEIN P28-SIS.
GN V-STs.
OS Simian sarcoma virus.
OC Viruses; Retroviridae; Mammalian type C retroviruses.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 83144004.
RA Devare S.G., Reddy E.P., Law J.D., Robbins K.C., Aaronson S.A.;
RT "Nucleotide sequence of the simian sarcoma virus genome:
RT demonstration that its acquired cellular sequences encode the
RT transforming gene product p28sis";
RL Proc. Natl. Acad. Sci. U.S.A. 80:731-735(1983).
CC -!- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; V01201; CAA24516.1; ALT_INIT.
DR PIR; A01381; TVMVSS.
DR HSP; P01127; 1PDG.
DR PFAM; PF00341; PDGF; 1.
DR PRINTS; PR00438; GFCVSKNOT.
DR PROSITE; PS00249; PDGF_1; 1.
KW Transforming protein; Oncogene; Growth factor.
SQ SEQUENCE 226 AA; 25411 MW; A16813ABB95B90C5 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 226;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 80 ABEKTR 85
|||||
QY 71 ABEKTR 76

RESULT 27
ID FANE.ECOLI STANDARD; PRT; 228 AA.
AC P25402;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CHAPERONE PROTEIN FANE PRECURSOR.

GN FANE.
OS Escherichia coli.
OG Plasmid pFK99.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ISOLATE B41;
RX MEDLINE; 91312125.
RA Bakker D., Vader C.E.M., Roosendaal B., Mooi F.R., Oudega B.,
RA de Graaf F.K.;
RT "Structure and function of periplasmic chaperone-like proteins
RT involved in the biosynthesis of K88 and K99 fimbriae in
RT enterotoxigenic Escherichia coli.";
RL Mol. Microbiol. 5:875-886(1991).
RN [2]
RP SEQUENCE OF 204-228 FROM N.A.
RX MEDLINE; 94187244.
RA Abe N., Moriishi K., Saito M., Naki M.;
RT "Confirmed nucleotide sequence of fanF of Escherichia coli K99
RT fimbriae";
RL Jpn. J. Vet. Res. 41:97-99(1993).
RN [3]
RP SEQUENCE OF 207-228 FROM N.A.
RC STRAIN-ISOLATE B41;
RX MEDLINE; 91211613.
RA Simons B.L., Willemssen P.T.J., Bakker D., Roosendaal B.,
RA de Graaf F.K., Oudega B.;
RT "Structure, localization and function of FanF, a minor component of
RT K99 fimbriae of enterotoxigenic Escherichia coli.";
RL Mol. Microbiol. 4:2041-2050(1990).
CC -!- FUNCTION: MEDIATES ASSEMBLY OF PILI BY FORMING SOLUBLE MULTIMERIC
CC COMPLEXES WITH PILI SUBUNITS AS AN INTERMEDIATE STEP IN THE
CC -!- SUBCELLULAR LOCATION: PERIPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE PERIPLASMIC PILUS CHAPERONE FAMILY.
CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X56001; CAA39474.1; -.
DR EMBL; S70131; AAB30305.1; -.
DR PIR; S12391; S12391.
DR PFAM; PF00345; pili-assembly; 1.
DR PRINTS; PR00969; CHAPERONPILI.
DR PROSITE; PS00835; PILI_CHAPERONE; 1.
KW Chaperone; Fimbria; Periplasmic; Signal; Immunoglobulin domain;
KW Plasmid.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 228 CHAPERONE PROTEIN FANE.
FT DISULFID 157 198 POTENTIAL.
SQ SEQUENCE 228 AA; 25314 MW; 5F1666C15EEFB949 CRC64;
Query Match 5.9%; Score 6; DB 1; Length 228;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 177 SIDLNS 182
|||||
QY 33 SIDLNS 38

RESULT 28
ID NARW.ECOLI STANDARD; PRT; 231 AA.
AC P19317;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)

```

DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE RESPIRATORY NITRATE REDUCTASE 2 DELTA CHAIN (EC 1.7.99.4).
GN NARW.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91042410.
RA Blasco F., Iobbi C., Ratouchniak J., Bonnefoy V., Chippaux M.;
RT "Nitrate reductases of Escherichia coli: sequence of the second
PT nitrate reductase and comparison with that encoded by the narGHJI
RL operon.";
RN Mol. Gen. Genet. 222:104-111(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN-K12 / MG1655;
RX MEDLINE; 97426617.
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner F.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RN Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN-K12;
RX MEDLINE; 97251357.
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
RA Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,
RA Motomura K., Nakade S., Nakamura Y., Nashimoto H., Nishio Y.,
RA Oshima T., Saito N., Sempel G., Seki Y., Sivasubram S.,
RA Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,
RA Yamamoto Y., Horiuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RN corresponding to the 28.0-40.1 min region on the linkage map.";
RN DNA Res. 3:363-377(1996).
RL -1- FUNCTION: REQUIRED FOR THE ASSEMBLY OF THE NITRATE REDUCTASE-
CC CYTOCHROME B-NR COMPLEX TO BE FULLY ACTIVE IN THE MEMBRANE
CC (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: NITRITE + ACCEPTOR = NITRATE + REDUCED
CC ACCEPTOR.
CC -1- SIMILARITY: 69% IDENTITY TO NARJ (THE DELTA CHAIN OF THE FIRST
CC E. COLI NITRATE REDUCTASE ENZYME).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X17110; CAA34966.1; -.
CC DR EMBL; AE000243; AAC74548.1; -.
CC DR EMBL; D90786; CAB21022.1; -.
CC DR EMBL; D90787; CAB21036.1; -.
CC DR PIR; S11429; S11429.
CC DR ECGENE; EG10645; NARW
KW Nitrate assimilation; Oxidoreductase.
SQ SEQUENCE 231 AA; 26160 MW; 30AE8EC9AE6290AB CRC64;

Query Match 5.9%; Score 6; DB 1; Length 231;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 156 QLAGSS 161
Qy 62 QLAGSS 67

RESULT 29

```

```

ID PDGB_SHEEP STANDARD; PRT; 241 AA.
AC Q95229;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PLATELET-DERIVED GROWTH FACTOR, B CHAIN PRECURSOR (PDGF B-CHAIN)
DE (PDGF-2).
GN PDGFB.
OS Eukaryotes (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-TEXEL; TISSUE-SPLEEN;
RX Woodall C.J., Zhang Z., Watt N.J.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PLATELET-DERIVED GROWTH FACTOR IS A POTENT MITOGEN FOR
CC CELLS OF MESENCHYMAL ORIGIN. BINDING OF THIS GROWTH FACTOR TO ITS
CC AFFINITY RECEPTOR ELICITS A VARIETY OF CELLULAR RESPONSES. IT IS
CC RELEASED BY PLATELETS UPON WOUNDING AND PLAYS AN IMPORTANT ROLE
CC IN STIMULATING ADJACENT CELLS TO GROW AND THEREBY HEAL THE WOUND.
CC -1- SUBUNIT: ANTIPARALLEL DISULFIDE-LINKED DIMER OF NONIDENTICAL (A
CC AND B) CHAINS. HOMODIMERS OF A AND B CHAINS ARE IMPLICATED IN
CC TRANSFORMATION PROCESSES.
CC -1- MISCELLANEOUS: A-A AND B-B, AS WELL AS A-B, DIMERS CAN BIND TO THE
CC PDGF RECEPTOR.
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X97123; CAA65790.1; -.
CC DR HSSP; P01127; 1PDG.
CC DR PFAM; PF00341; PDGF.
CC DR PRINTS; PR00438; GFCYSKNOT.
CC DR PROSITE; PS00249; PDGF_1; 1.
KW Mitogen; Growth factor; Proto-oncogene; Platelet; Signal.
FT SIGNAL 1 20
FT PROPEP 21 81
FT CHAIN 82 190
FT PROPEP 191 241
FT SITE 108 108
FT SITE 111 111
FT DISULFID 97 141
FT DISULFID 130 178
FT DISULFID 134 180
FT DISULFID 124 124
FT DISULFID 133 133
FT CARBOHYD 63 63
SQ SEQUENCE 241 AA; 27331 MW; 37BE1EC12E7D2863 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 241;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 95 AECKTR 100
Qy 71 AECKTR 76

RESULT 30
ID PDGB_MOUSE STANDARD; PRT; 241 AA.
AC P31240;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE PLATELET-DERIVED GROWTH FACTOR, B CHAIN PRECURSOR (PDGF B-CHAIN)

```


[9] SEQUENCE OF 26-241 FROM N.A.
 RX MEDLINE; 86194981.
 RA Weich H.A., Sebald W., Schairer H.U., Hoppe J.;
 RT "The human osteosarcoma cell line U-2 OS expresses a 3.8 kilobase
 RL mRNA which codes for the sequence of the PDGF-B chain.";
 RN FEBS Lett. 198;344-348(1986).
 [10]
 RN SEQUENCE OF 153-200 FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE; 84236121.
 RA Johansson A., Heldin C.H., Wasteson A., Westermark B., Deuel T.F.,
 RA Huang J.S., Seeburg P.H., Gray A., Ullrich A., Strack G.,
 RA Stroobant P., Waterfield M.D.;
 RT "The c-sis gene encodes a precursor of the B chain of
 RL platelet-derived growth factor.";
 RN EMBO J. 3:921-928(1984).
 [11]
 RN SEQUENCE OF 82-110.
 RX MEDLINE; 83197379.
 RA Antoniadou H.N., Hunkapiller M.W.;
 RT "Human platelet-derived growth factor (PDGF): amino-terminal amino
 RL acid sequence.";
 RN Science 220:963-965(1983).
 [12]
 RN SEQUENCE OF 82-112.
 RX MEDLINE; 83244981.
 RA Waterfield M.D., Strack G.T., Whittle N., Stroobant P., Johansson A.,
 RA Wasteson A., Westermark B., Heldin C.H., Huang J.S., Deuel T.F.;
 RT "Platelet-derived growth factor is structurally related to the
 RL putative transforming protein p28sis of simian sarcoma virus.";
 RN Nature 304:35-39(1983).
 [13]
 RN MUTAGENESIS, & IMPORTANCE OF ARG-108 AND ILE-111 FOR RECEPTOR-BINDING.
 RX MEDLINE; 92097530.
 RA Clements J.M., Bowden L.J., Bloxidge R.E., Catlin G., Cook A.L.,
 RA Craig S., Drummond A.H., Edwards R.M., Fallon A., Green D.R.,
 RA Hellewell P.G., Kirwin P.M., Nacey P.D., Richardson S.J., Brown D.,
 RA Chahwa S.B., Snarey M., Winslow D.;
 RT "Two PDGF-B chain residues, arginine 27 and isoleucine 30, mediate
 RL receptor binding and activation.";
 RN EMBO J. 10:4113-4120(1991).
 [14]
 RN INTERCHAIN DISULFIDE BONDS.
 RX MEDLINE; 92283833.
 RA Andersson M., Oestman A., Baekstroem G., Hellman U.,
 RA George-Nascimento C., Westermark B., Heldin C.H.;
 RT "Assignment of interchain disulfide bonds in platelet-derived growth
 RL factor (PDGF) and evidence for agonist activity of monomeric PDGF.";
 RN J. Biol. Chem. 267:11260-11266(1992).
 [15]
 RN X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
 RX MEDLINE; 93010987.
 RA Osner C., D'Arcy A., Winkler F.K., Eggmann B., Hosang M.;
 RT "Crystal structure of human platelet-derived growth factor BB.";
 RN EMBO J. 11:3921-3926(1992).
 CC -1- FUNCTION: PLATELET-DERIVED GROWTH FACTOR IS A POTENT MITOGEN FOR
 CC CELLS OF MESENCHYMAL ORIGIN. BINDING OF THIS GROWTH FACTOR TO ITS
 CC AFFINITY RECEPTOR ELICITS A VARIETY OF CELLULAR RESPONSES. IT IS
 CC RELEASED BY PLATELETS UPON WOUNDING AND PLAYS AN IMPORTANT ROLE
 CC IN STIMULATING ADJACENT CELLS TO GROW AND THEREBY HEAL THE WOUND.
 CC -1- SUBUNIT: ANTIPARALLEL DISULFIDE-LINKED DIMER OF NONIDENTICAL (A
 CC AND B) CHAINS. HOMODIMERS OF A AND B CHAINS ARE IMPLICATED IN
 CC TRANSFORMATION PROCESSES.
 CC -1- PHARMACEUTICAL: AVAILABLE UNDER THE NAME REGRANEX (ORTHO-MCNEIL).
 CC USED TO PROMOTE HEALING IN DIABETIC NEUROPATHIC FOOT ULCERS.
 CC -1- MISCELLANEOUS: A-A AND B-B, AS WELL AS A-B, DIMERS CAN BIND TO THE
 CC PDGF RECEPTOR.
 CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
 CC -1- DATABASE: NAME-RED systems' cytokine source book;
 CC WWW="http://www.rndsystems.com/cyt_cat/pdgf.html".
 CC -1- DATABASE: NAME-Regranex; NOTE-Clinical information on Regranex;
 CC WWW="http://www.regranex.com/".

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; K01401; AAA60552.1; -
 DR EMBL; K01918; AAA60552.1; JOINED.
 DR EMBL; J00121; AAA60552.1; JOINED.
 DR EMBL; K01398; AAA60552.1; JOINED.
 DR EMBL; K01399; AAA60552.1; JOINED.
 DR EMBL; K01400; AAA60552.1; JOINED.
 DR EMBL; X02811; CAA26579.1; -
 DR EMBL; M12783; AAA60553.1; -
 DR EMBL; X02744; CAA26524.1; -
 DR EMBL; K01917; AAA98793.1; -
 DR EMBL; K01913; AAA98793.1; JOINED.
 DR EMBL; K01914; AAA98793.1; JOINED.
 DR EMBL; K01915; AAA98793.1; JOINED.
 DR EMBL; K01916; AAA98793.1; JOINED.
 DR EMBL; X03702; CAA27333.1; -
 DR EMBL; Z81010; CAB02635.1; -
 DR EMBL; X00561; CAA25228.1; -
 DR EMBL; X00561; CAA25229.1; -
 DR EMBL; X98706; CAA67262.1; -
 DR PIR; A94276; PFHUG2.
 DR PDB; LPDG; 31-JAN-94.
 DR MIM; L90040; -
 DR PFAM; PF00341; PDGF; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PROSITE; PS00249; PDGF_1; 1.
 KW Mitogen; Growth factor; Proto-oncogene; Platelet; Signal;
 KW Pharmaceutical; 3D-structure.
 FT SIGNAL 1 20
 FT PROPEP 21 81
 FT CHAIN 82 190
 FT PROPEP 191 241
 FT SITE 108 108
 FT SITE 111 111
 FT DISULFID 97 141
 FT DISULFID 130 178
 FT DISULFID 134 180
 FT DISULFID 124 124
 FT DISULFID 133 133
 FT CONFLICT 21 21
 FT CONFLICT 101 101
 FT CONFLICT 105 105
 FT CONFLICT 107 107
 FT STRAND 90 91
 FT STRAND 94 94
 FT STRAND 98 103
 FT STRAND 120 121
 FT STRAND 124 131
 FT TURN 137 138
 ... Note: remainder of annotations omitted.
 Query Match 5.9%; Score 6; DB 1; Length 241;
 Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 95 AECKTR 100
 QY 71 AECKTR 76
 RESULT 32
 ID PDGF_FELCA STANDARD; PRT; 245 AA.
 AC P12919;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PLATELET-DERIVED GROWTH FACTOR, B CHAIN PRECURSOR (PDGF B-CHAIN)
DE (PDGFB) (PDGFB-2).
GN PDGFB OR C-SIS.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 87146463.
RA van den Ouweland A.M.W., van Groningen J.J.M., Schalken J.A.,
RA van Neck H.W., Bloemers H.P.J., van de Ven W.J.M.;
RT "Genetic organization of the c-sis transcription unit.";
RL Nucleic Acids Res. 15:959-970(1987).
RN [2]
RP REVISIONS.
RA van den Ouweland A.M.W.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PLATELET-DERIVED GROWTH FACTOR IS A POTENT MITOGEN FOR
CC CELLS OF MESENCHYMAL ORIGIN. BINDING OF THIS GROWTH FACTOR TO ITS
CC AFFINITY RECEPTOR ELICITS A VARIETY OF CELLULAR RESPONSES. IT IS
CC RELEASED BY PLATELETS UPON WOUNDING AND PLAYS AN IMPORTANT ROLE
CC IN STIMULATING ADJACENT CELLS TO GROW AND THEREBY HEAL THE WOUND.
CC -!- SUBUNIT: ANTIPARALLEL DISULFIDE-LINKED DIMER OF NONIDENTICAL (A
CC AND B) CHAINS. HOMODIMERS OF A AND B CHAINS ARE IMPLICATED IN
CC TRANSFORMATION PROCESSES.
CC -!- MISCELLANEOUS: A-A AND B-B, AS WELL AS A-B, DIMERS CAN BIND TO THE
CC PDGF RECEPTOR.
CC -!- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X05112; CAA28758.1; ALT_SEQ.
DR PIR; A26402; TVCTSS.
DR HSSP; P01127; 1PDG.
DR PFAM; PF00341; PDGF_1.
DR PRINTS; PR00438; GFCYSKNOT.
DR PROSITE; PS00249; PDGF_1; 1.
KW Mitogen; Growth factor; Proto-oncogene; Platelet; Signal.
FT SIGNAL 1 20
FT PROPEP 21 81
FT CHAIN 82 194
FT PROPEP 195 245
FT DISULFID 101 145
FT DISULFID 134 182
FT DISULFID 138 184
FT DISULFID 128 128
FT DISULFID 137 137
FT INTERCHAIN (BY SIMILARITY).
FT INTERCHAIN (BY SIMILARITY).
SQ SEQUENCE 245 AA; 27787 MW; E771529ID9837512 CRC64;
Query Match 5.9%; Score 6; DB 1; Length 245;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 99 AECCTR 104
Qy 71 AECCTR 76
RESULT 33
ID VGL2_EBV STANDARD; PRT; 248 AA.
AC P03218;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE PROBABLE MEMBRANE GLYCOPROTEIN.
GN BILF2.

OS Epstein-barr virus (strain B95-8) (Human herpesvirus 4).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 85035713.
RA Bankier A.T., Deininger P.L., Farrell P.J., Barrell B.G.;
RT "Sequence analysis of the 17,166 base-pair EcoRI fragment C of B95-8
RT Epstein-Barr virus.";
RL Mol. Biol. Med. 1:21-45(1983).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 84270667.
RA Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
RA Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
RA Tuffnell P.S., Barrell B.G.;
RT "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
RL Nature 310:207-211(1984).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90244363.
RA Mackett M., Conway M.J., Arrand J.R., Haddad R.S.,
RA Hutt-Fletcher L.M.;
RT "Characterization and expression of a glycoprotein encoded by the
RT Epstein-Barr virus BamHI I fragment.";
RL J. Virol. 64:2545-2552(1990).
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; V01555; CAA24803.1; -.
DR EMBL; M37129; AAA45876.1; -.
DR PIR; A03780; QQBE4L.
DR PIR; S33052; S33052.
DR PFAM; PF00047; Ig_1.
KW Glycoprotein; Late protein; Membrane.
SQ SEQUENCE 248 AA; 27076 MW; C3F33A253B959ADA CRC64;
Query Match 5.9%; Score 6; DB 1; Length 248;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 78 TNGGYN 83
Qy 25 TNGGYN 30
RESULT 34
ID YKT0_CAEEL STANDARD; PRT; 254 AA.
AC P34321;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE HYPOTHETICAL 29.1 KDA PROTEIN C07A9.10 IN CHROMOSOME III.
GN C07A9.10.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE; 94150718.
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,

RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R.,
RA Sims M., Smaldon N., Smith A., Smith M., Sonnhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Watson R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RL elegans";
RL Nature 368:32-38(1994).
CC -1- SIMILARITY: TO C.ELEGANS F52C9.6 AND F23C11.1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z29094; CAA82337.1; -;
CC PIR; S40702; S40702.
CC WORMPEP; C07A9.10; CE00496.
CC PFAM; PF00442; UCH-1; 1.
CC Hypothetical protein.
SQ SEQUENCE 254 AA; 29082 MW; 8057BF17FE13CB50 CRC64;
Query Match 5.9%; Score 6; DB 1; Length 254;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 186 LAECK 191
| | | | |
Qy 69 LAECK 74
RESULT 35
ID RNPH-AQUAE STANDARD; PRT; 255 AA.
AC 067069;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE 15-FEB-2000 (Rel. 39, Last annotation update)
DE RIBONUCLEASE PH (EC 2.7.7.56) (RNASE PH) (TRNA
DE NUCLEOTIDYLTRANSFERASE).
GN RPH OR RNPH OR AQ-924.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN-VF5;
CC MEDLINE; 98196666.
CC Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
CC Graham D.E., Overbeek R., Snead M.A., Keller M., AuJay M., Huber R.,
CC Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
CC "The complete genome of the hyperthermophilic bacterium Aquifex
CC aeolicus";
CC Nature 392:353-358(1998).
CC -1- FUNCTION: RNASE PH IS A PHOSPHORYLYC EXORIBONUCLEASE THAT
CC REMOVES NUCLEOTIDE RESIDUES FOLLOWING THE -CCA TERMINUS OF TRNA
CC AND ADDS NUCLEOTIDES TO THE ENDS OF RNA MOLECULES BY USING
CC NUCLEOSIDE DIPHOSPHATES AS SUBSTRATES (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: TRNA(N+1) + ORTHOPHOSPHATE = TRNA(N) +
CC A NUCLEOSIDE DIPHOSPHATE.
CC -1- SIMILARITY: BELONGS TO THE RNASE PH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000714; AAC07032.1; -;
CC PFAM; PF01138; RNase_PH; 1.
DR

DR PROSITE; PS01277; RIBONUCLEASE_PH; 1.
KW Transferase; Nucleotidyltransferase; tRNA processing.
SQ SEQUENCE 255 AA; 28372 MW; 7131CCD8AEDC3917 CRC64;
Query Match 5.9%; Score 6; DB 1; Length 255;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 42 SVIENV 47
| | | | |
Qy 38 SVIENV 43
RESULT 36
ID MS2A_DROME STANDARD; PRT; 264 AA.
AC P1033;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MALE ACCESSORY GLAND SECRETORY PROTEIN 355A PRECURSOR.
GN ACP26AA OR MST26AA OR MST355A.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN-CANTON-S;
CC MEDLINE; 89053045.
CC Monsma S.A., Wolfner M.F.;
CC "Structure and expression of a Drosophila male accessory gland gene
CC whose product resembles a peptide pheromone precursor.";
CC Genes Dev. 2:1063-1073(1988).
CC [2]
CC SEQUENCE FROM N.A.
CC STRAIN-NC1, NC2, NC3, NC4, NC5, NC6, NC7, NC8, NC9 AND NC10;
CC MEDLINE; 93106377.
CC Aquade M., Miyashita N., Langley C.H.;
CC "Polymorphism and divergence in the Mst26A male accessory gland gene
CC region in Drosophila";
CC Genetics 132:753-770(1992).
CC -1- FUNCTION: THIS PROTEIN IS TRANSFERRED FROM MALE TO FEMALE'S
CC HEMOLymph DURING MATING, AFFECTING EGG-LAYING AND BEHAVIOR AFTER
CC MATING.
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- TISSUE SPECIFICITY: MAIN CELLS OF THE ACCESSORY GLANDS OF MALES.
CC -1- PTM: IT UNDERGOES SEVERAL CLEAVAGES AS IT IS SECRETED AND IT
CC IS FURTHER PROCESSED IN THE RECIPIENT FEMALE.
CC -1- SIMILARITY: REGION OF HOMOLOGY WITH APLYSIA CALIFORNICA EGG-LAYING
CC HORMONE (ELH).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y00219; CAA68366.1; -;
CC EMBL; X70888; CAA50232.1; -;
CC EMBL; X70889; CAA50234.1; -;
CC EMBL; X70890; CAA50236.1; -;
CC EMBL; X70891; CAA50238.1; -;
CC EMBL; X70892; CAA50240.1; -;
CC EMBL; X70893; CAA50242.1; -;
CC EMBL; X70894; CAA50244.1; -;
CC EMBL; X70895; CAA50246.1; -;
CC EMBL; X70896; CAA50248.1; -;
CC EMBL; X70897; E64715; ALT_FRAME.
CC PIR; S02853; S02853.
CC PIR; S30409; S30409.
CC PIR; S30411; S30411.
DR

DR PIR: S30413; S30413.
DR PIR: S30415; S30415.
DR PIR: S30423; S30423.
DR PIR: S30425; S30425.
DR FLYBASE: FBgn0002855; Acp26Aa.
KW Glycoprotein; Behavior; Signal.
FT SIGNAL 1 18
FT CHAIN 19 264
FT
FT CARBOHYD 88 88
FT CARBOHYD 122 122
FT CARBOHYD 138 138
FT CARBOHYD 145 145
FT CARBOHYD 150 150
FT SIMILAR 120 137
FT VARIANT 19 19
FT VARIANT 24 24
FT VARIANT 32 32
FT VARIANT 46 46
FT VARIANT 76 76
FT VARIANT 79 79
FT VARIANT 101 101
FT VARIANT 109 109
FT VARIANT 207 207
FT VARIANT 221 221
SQ SEQUENCE 264 AA; 29671 MW; BD64298CE04DFDE5 CRC64;
Query Match 5.9%; Score 6; DB 1; Length 264;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 96 SSIDLN 101
Qy 32 SSIDLN 37
RESULT 37;
ID YAIM_HAEIN STANDARD; PRT; 275 AA.
AC P44556;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE HYPOTHEICAL PROTEIN H1084.
GN H10184.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
CC Haemophilus.
CC [1]
RP SEQUENCE FROM N.A.
RX STRAIN-RD / KW20;
RX MEDLINE; 95350630.
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Georghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.";
RL Science 269:496-512(1995).
CC -!- SIMILARITY: STRONG, TO E.COLI YAIM AND YEIG.
CC -!- SIMILARITY: STRONG, TO HUMAN ESTERASE D.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch.

CC or send an email to license@isb-sib.ch.
CC EMBL; U32703; AAC21853.1; -.
DR TIGR; H10184; -.
DR PFAM; PF00756; Esterase; 1.
KW Hypothetical protein.
SQ SEQUENCE 275 AA; 31317 MW; 781F5C0411546D3D CRC64;
Query Match 5.9%; Score 6; DB 1; Length 275;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 233 FIETCR 238
Qy 54 FIETCR 59
RESULT 38
ID HEM3_CHLVI STANDARD; PRT; 279 AA.
AC P28464;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PORPHOBILINOMEN DEAMINASE (EC 4.3.1.8) (PBG) (HYDROXYMETHYLBILANE
DE SYNTHASE) (HMBS) (PRE-UROPORPHYRINOGEN SYNTHASE).
GN HEMC.
OS Chlorobium vibrioforme.
OC Bacteria; Green sulfur bacteria; Chlorobium.
CC [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F. THIOSULFATOPHILUM / NCIB 8327;
RA Majumdar D., Wyche J.H.;
RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-43 FROM N.A.
RX STRAIN=F. THIOSULFATOPHILUM / NCIB 8327;
RX MEDLINE; 92171712.
RA Majumdar D., Avissar Y.J., Wyche J.H., Beale S.I.;
RT "Structure and expression of the Chlorobium vibrioforme hema gene.";
RL Arch. Microbiol. 156:281-289(1991).
CC -!- FUNCTION: TETRAPOLYMERIZATION OF THE MONOPYRROLE PBG INTO THE
CC HYDROXYMETHYLBILANE PREUROPORPHYRINOGEN IN SEVERAL DISCRETE STEPS.
CC -!- CATALYTIC ACTIVITY: 4 PORPHOBILINOGEN + 4 NH(3) =
CC HYDROXYMETHYLBILANE + 4 H(2)O =
CC -!- COFACTOR: COVALENTLY BINDS A DIPYRROMETHANE COFACTOR TO WHICH
CC THE PORPHOBILINOGEN SUBUNITS ARE ADDED.
CC -!- PATHWAY: FOURTH STEP IN PORPHYRIN BIOSYNTHESIS BY THE C5 PATHWAY.
CC INVOLVED IN CHLOROPHYLL BIOSYNTHESIS.
CC -!- SUBUNIT: MONOMER.
CC -!- SIMILARITY: BELONGS TO THE HMBS FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch.

Query Match 5.9%; Score 6; DB 1; Length 279;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 219 IDGTLK 224

QY 94 IDGTLK 99

RESULT 39

ID YLMW3 CAEEL STANDARD; PRT; 288 AA.
AC P34412; 1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 29, Last annotation update)
DE HYPOTHETICAL 32.9 KDA PROTEIN F22B7.9 IN CHROMOSOME III.
GN F22B7.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-BRISTOL N2;
RX MEDLINE; 94150718.
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkneen R.,
RA Sims M., Smaldon N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterson R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohlman P.;
FT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans.";
RL Nature 368:32-38(1994).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; L12018; AAA65465.1; -

DR WORKPEP; F22B7.9; CE00162.

KW Hypothetical protein.

SEQUENCE 288 AA; 32856 MW; 2E7583B33DD23F8C CRC64;

Query Match 5.9%; Score 6; DB 1; Length 288;

Best Local Similarity 100.0%; Pred. No. 8.75e+00;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 199 NVDGSL 204

QY 42 NVDGSL 47

RESULT 40

ID PYR3_PREDI STANDARD; PRT; 288 AA.
AC P11400;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE PHYCOBILISOME 39 KDA LINKER POLYPEPTIDE, PHYCOCYANIN-ASSOCIATED, ROD
DE (L-39/R).
GN CPC12.
OS Fremyella diplosiphon (Calothrix PCC 7601).
OC Bacteria; Cyanobacteria; Nostocales; Rivulariaceae; Fremyella.
[1]
RN SEQUENCE FROM N.A.
RP PFAM; PF00318; Ribosomal_S2; 1.
RX MEDLINE; 87222193.

RA Lomax T.L., Conley P.B., Schilling J., Grossman A.R.;
RT "Isolation and characterization of light-regulated phycobilisome
RT linker polypeptide genes and their transcription as a polycistronic
RT mRNA.";
RL J. Bacteriol. 169:2675-2684(1987).
CC -1- FUNCTION: ROD LINKER PROTEIN, ASSOCIATED WITH PHYCOCYANIN. LINKER
CC POLYPEPTIDES DETERMINE THE STATE OF AGGREGATION AND THE LOCATION
CC OF THE DISC-SHAPED PHYCOBILIPROTEIN UNITS WITHIN THE PHYCOBILISOME
CC AND MODULATE THEIR SPECTROSCOPIC PROPERTIES IN ORDER TO MEDIATE A
CC DIRECTED AND OPTIMAL ENERGY TRANSFER.
CC -1- SUBCELLULAR LOCATION: THIS PROTEIN OCCURS IN THE ROD, IT IS
CC ASSOCIATED WITH PHYCOCYANIN.
CC -1- SIMILARITY: TO OTHER PHYCOBILISOME LINKER PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; M16490; AAA24887.1; -

DR PIR; B25974; B25974.

DR PFAM; PF01383; CpcB; 1.

DR PFAM; PF00427; PBS_Linker_poly; 1.

KW Phycobilisome; Photosynthesis.

FT INIT_MET 0

SQ SEQUENCE 288 AA; 32184 MW; D5A325E06A0B2F6 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 288;

Best Local Similarity 100.0%; Pred. No. 8.75e+00;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 181 QLAGSS 186

QY 62 QLAGSS 67

RESULT 41

ID RS2_RICPR STANDARD; PRT; 296 AA.
AC Q3ZE61;
DT 15-FEB-2000 (Rel. 39, Created)
DT 15-FEB-2000 (Rel. 39, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE 30S RIBOSOMAL PROTEIN S2.
GN RPSB OR RP086.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=MADRID E;
RX MEDLINE; 99039499.
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sicheritz-Ponten T., Almark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria.";
RL Nature 396:133-140(1998).
CC -1- SIMILARITY: BELONGS TO THE S2P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AJ235270; CAAL4556.1; -
DR PFAM; PF00318; Ribosomal_S2; 1.
DR PROSITE; PS00962; RIBOSOMAL_S2_1; 1.


```

DR PROSITE; PS00963; RIBOSOMAL_S2_2; 1.
KW Ribosomal protein.
SQ SEQUENCE 296 AA; 32961 MW; 386B925D43BD6C0 CRC64;

Query Match
  5.9%; Score 6; DB 1; Length 296;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 72 FVSTKI 77
   |||||
Qy 80 FVSTKI 85

RESULT 42
ID SFMH_ECOLI STANDARD; PRT; 325 AA.
AC P75715; P77078;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE SFMH PROTEIN PRECURSOR.
GN SFMH.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE; 97426617.
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner F.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12.";
RT Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Chung E., Allen E., Araujo R., Aparicio A., Davis K., Duncan M.,
RA Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O., Lew H., Lin D.,
RA Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INVOLVED IN REGULATION OF LENGTH AND MEDIATION OF
CC ADHESION OF SFMA FIMBRIAE (BUT NOT NECESSARY FOR THE PRODUCTION
CC OF FIMBRIAE). SEEMS TO BE MANNOSE BINDING ADHESIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE FIMH / LPFD FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000159; AAC73635.1; -.
DR EMBL; U82598; AAB40731.1; ALT_INIT.
DR ECOGENE; EG13884; SFMH.
KW Fimbria; Signal.
FT SIGNAL 1 ? POTENTIAL.
FT CHAIN 1 ? 325 SFMH PROTEIN.
SQ SEQUENCE 325 AA; 35493 MW; B607915A95542193 CRC64;

Query Match
  5.9%; Score 6; DB 1; Length 325;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 278 DLNSVI 283
   |||||
Qy 35 DLNSVI 40

RESULT 43
ID RLUD_BORBU STANDARD; PRT; 326 AA.
AC P70870; O51050;

```

```

DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE RIBOSOMAL LARGE SUBUNIT PSEUDOURIDINE SYNTHASE D (EC 4.2.1.70)
DE (PSEUDOURIDYLATE SYNTHASE) (URACIL HYDROLASE).
GN RLUD OR BB0018.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HB19;
RA Boursaux-Eude C., Margarita D., Belfaiza J., Old I.G.,
RA Saint-Girons I.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE; 98065943.
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi.";
RL Nature 390:580-586(1997).
CC -!- FUNCTION: RESPONSIBLE FOR SYNTHESIS OF PSEUDOURIDINE FROM URACIL
CC AT TWO POSITIONS IN 23S RIBOSOMAL RNA (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: URACIL + D-RIBOSE 5-PHOSPHATE = PSEUDOURIDINE
CC 5'-PHOSPHATE + H(2)O.
CC -!- SIMILARITY: BELONGS TO THE RLU FAMILY OF PSEUDOURIDINE SYNTHASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y09142; CAA70352.1; -.
DR EMBL; AE001116; AAC66413.1; ALT_INIT.
DR TIGR; BB0018; -.
DR PFAM; PF00849; Pseudou synth_2; 1.
DR PROSITE; PS01129; PSL_RLU; 1.
KW Lyase.
FT CONFLICT 172 172 T -> I (IN REF. 1).
FT CONFLICT 268 268 K -> E (IN REF. 1).
FT CONFLICT 299 326 FVFLSNFYKSDLNLIIDNLVFLRDF ->
FT ICYFFVFLQ (IN REF. 1).
SQ SEQUENCE 326 AA; 38067 MW; 52266C0AF8BEA3C CRC64;

Query Match
  5.9%; Score 6; DB 1; Length 326;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 263 VDGLSK 268
   |||||
Qy 43 VDGLSK 48

RESULT 44
ID VGH_BPPHX STANDARD; PRT; 328 AA.
AC P03646;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MINOR SPIKE PROTEIN (H PROTEIN) (PILOT PROTEIN).
GN H.
OS Bacteriophage phi-X174.

```

OC Viruses; ssDNA viruses; Microviridae; Microvirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 77171175.
RA Sanger F., Air G.M., Barrell B.G., Brown N.L., Coulson A.R.,
RA Fides J.C., Hutchison C.A. III, Slocombe P.M., Smith M.;
RA "Nucleotide sequence of bacteriophage phi X174 DNA.";
RT Nature 265:687-695(1977).
RL [2]
RN REVISIONS.
RP MEDLINE; 79091185.
RX Sanger F., Coulson A.R., Friedmann T., Air G.M., Barrell B.G.,
RA Brown N.L., Fiddes J.C., Hutchison C.A. III, Slocombe P.M., Smith M.;
RA "The nucleotide sequence of bacteriophage phiX174.";
RT J. Mol. Biol. 125:325-246(1978).
RL [3]
RN X-RAY CRYSTALLOGRAPHY.
RP MEDLINE; 92107297.
RX McKenna R., Xia D., Williangmann P., Ilag L.L., Krishnaswamy S.,
RA Rossmann M.G., Olson N.H., Baker T.S., Incardona N.L.;
RA "Atomic structure of single-stranded DNA bacteriophage phi X174 and
RA its functional implications.";
RP Nature 355:137-143(1992).
RL [4]
RN X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RP MEDLINE; 94210479.
RX McKenna R., Ilag L.L., Rossmann M.G.;
RA "Analysis of the single-stranded DNA bacteriophage phi X174, refined
RA at a resolution of 3.0 A.";
RT J. Mol. Biol. 237:517-543(1994).
RL [5]
RN -1- FUNCTION: MINOR SPIKE COMPONENT OF THE VIRAL SHELL. H PROTEIN IS
CC INVOLVED IN THE EJECTION OF THE PHAGE DNA IN THE HOST AND IS
CC INJECTED WITH THE DNA IN THE PERIPLASMIC SPACE OF THE HOST.
CC -1- SUBUNIT: THE VIRION IS COMPOSED OF 60 COPIES EACH OF THE F, G, AND
CC J PROTEINS, AND 12 COPIES OF THE H PROTEIN. THERE ARE 12 SPIKES
CC WHICH ARE EACH COMPOSED OF 5 G AND ONE H PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; V01128; CAA24360.1; -;
DR EMBL; J02482; AAA32580.1; -;
DR PIR; A04254; ZHBPF4.
RW Coat protein.
SQ SEQUENCE 328 AA; 34419 MW; 3793DD830678927C CRC64;
Query Match 5.9%; Score 6; DB 1; Length 328;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 54 SAIQGS 59
QY 11 SAIQGS 16
RESULT 45
ID VGH_BPS13 STANDARD; PRT; 328 AA.
AC P07933;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DE 01-OCR-1996 (Rel. 34, Last annotation update)
DE MINOR SPIKE PROTEIN (H PROTEIN) (PILOT PROTEIN).
GN H.
OS Bacteriophage S13.
OC Viruses; ssDNA viruses; Microviridae; Microvirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 86165869.
RA Sanger F., Air G.M., Barrell B.G., Brown N.L., Coulson A.R.,
RA Fides J.C., Hutchison C.A. III, Slocombe P.M., Smith M.;
RA "Nucleotide sequence of bacteriophage phi X174 DNA.";
RT Nature 265:687-695(1977).
RL [2]
RN REVISIONS.
RP MEDLINE; 79091185.
RX Sanger F., Coulson A.R., Friedmann T., Air G.M., Barrell B.G.,
RA Brown N.L., Fiddes J.C., Hutchison C.A. III, Slocombe P.M., Smith M.;
RA "The nucleotide sequence of bacteriophage phiX174.";
RT J. Mol. Biol. 125:325-246(1978).
RL [3]
RN X-RAY CRYSTALLOGRAPHY.
RP MEDLINE; 92107297.
RX McKenna R., Xia D., Williangmann P., Ilag L.L., Krishnaswamy S.,
RA Rossmann M.G., Olson N.H., Baker T.S., Incardona N.L.;
RA "Atomic structure of single-stranded DNA bacteriophage phi X174 and
RA its functional implications.";
RP Nature 355:137-143(1992).
RL [4]
RN X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RP MEDLINE; 94210479.
RX McKenna R., Ilag L.L., Rossmann M.G.;
RA "Analysis of the single-stranded DNA bacteriophage phi X174, refined
RA at a resolution of 3.0 A.";
RT J. Mol. Biol. 237:517-543(1994).
RL [5]
RN -1- FUNCTION: MINOR SPIKE COMPONENT OF THE VIRAL SHELL. H PROTEIN IS
CC INVOLVED IN THE EJECTION OF THE PHAGE DNA IN THE HOST AND IS
CC INJECTED WITH THE DNA IN THE PERIPLASMIC SPACE OF THE HOST.
CC -1- SUBUNIT: THE VIRION IS COMPOSED OF 60 COPIES EACH OF THE F, G, AND
CC J PROTEINS, AND 12 COPIES OF THE H PROTEIN. THERE ARE 12 SPIKES
CC WHICH ARE EACH COMPOSED OF 5 G AND ONE H PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; V01128; CAA24360.1; -;
DR EMBL; J02482; AAA32580.1; -;
DR PIR; A04254; ZHBPF4.
RW Coat protein.
SQ SEQUENCE 328 AA; 34419 MW; 3793DD830678927C CRC64;
Query Match 5.9%; Score 6; DB 1; Length 328;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 54 SAIQGS 59
QY 11 SAIQGS 16
RESULT 45
ID VGH_BPS13 STANDARD; PRT; 328 AA.
AC P07933;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DE 01-OCR-1996 (Rel. 34, Last annotation update)
DE MINOR SPIKE PROTEIN (H PROTEIN) (PILOT PROTEIN).
GN H.
OS Bacteriophage S13.
OC Viruses; ssDNA viruses; Microviridae; Microvirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 86165869.

Lau P.C.K., Spencer J.H.;
RA "Nucleotide sequence and genome organization of bacteriophage S13
RA DNA.";
RL Gene 40:273-284(1985).
CC -1- FUNCTION: MINOR SPIKE COMPONENT OF THE VIRAL SHELL. H PROTEIN IS
CC INVOLVED IN THE EJECTION OF THE PHAGE DNA IN THE HOST AND IS
CC INJECTED WITH THE DNA IN THE PERIPLASMIC SPACE OF THE HOST.
CC -1- SUBUNIT: THE VIRION IS COMPOSED OF 60 COPIES EACH OF THE F, G, AND
CC J PROTEINS, AND 12 COPIES OF THE H PROTEIN. THERE ARE 12 SPIKES
CC WHICH ARE EACH COMPOSED OF 5 G AND ONE H PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M14428; AAA32592.1; -;
DR PIR; JS0459; JS0459.
RW Coat protein.
SQ SEQUENCE 328 AA; 34405 MW; EEABE8053765177F CRC64;
Query Match 5.9%; Score 6; DB 1; Length 328;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 54 SAIQGS 59
QY 11 SAIQGS 16
RESULT 46
ID VGH_BPAL3 STANDARD; PRT; 330 AA.
AC P03650;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-OCR-1996 (Rel. 34, Last annotation update)
DE MINOR SPIKE PROTEIN (H PROTEIN) (PILOT PROTEIN).
GN H.
OS Bacteriophage alpha-3.
OC Viruses; ssDNA viruses; Microviridae; Microvirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 92223109.
RA Kodaira K.-I., Nakano K., Okada S., Taketo A.;
RT "Nucleotide sequence of the genome of the bacteriophage alpha 3:
RT interrelationship of the genome structure and the gene products with
RT those of the phages, phi X174, G4 and phi K.";
RL Biochim. Biophys. Acta 1130:277-288(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 85226468.
RA Kodaira K.-I., Nakano K., Taketo A.;
RT "Function and structure of microvirid phage alpha 3 genome. DNA
RT sequence of H gene and properties of missense H mutant.";
RL Biochim. Biophys. Acta 825:255-260(1985).
RN [3]
RP SEQUENCE OF 1-70 FROM N.A.
RX MEDLINE; 80049950.
RA Sims J., Capon D., Dressler D.;
RT "dnaG (primase)-dependent origins of DNA replication. Nucleotide
RT sequences of the negative strand initiation sites of bacteriophages
RT St-1, phi K, and alpha 3.";
RL J. Biol. Chem. 254:12615-12628(1979).
CC -1- FUNCTION: MINOR SPIKE COMPONENT OF THE VIRAL SHELL. H PROTEIN IS
CC INVOLVED IN THE EJECTION OF THE PHAGE DNA IN THE HOST AND IS
CC INJECTED WITH THE DNA IN THE PERIPLASMIC SPACE OF THE HOST.
CC -1- SUBUNIT: THE VIRION IS COMPOSED OF 60 COPIES EACH OF THE F, G, AND
CC J PROTEINS, AND 12 COPIES OF THE H PROTEIN. THERE ARE 12 SPIKES
CC WHICH ARE EACH COMPOSED OF 5 G AND ONE H PROTEINS.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X60322; CAA42883.1; -
 DR EMBL; M25640; AAA32174.1; -
 DR EMBL; J02444; AAA32176.1; -
 DR PIR; A04257; ZHBPA3.
 DR PIR; A21537; A21537.
 DR PIR; S22332; S22332.
 KW Coat protein.
 SQ SEQUENCE 330 AA; 34844 MW; B194EF4A83CE3CE8 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 330;
 Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 58 SAIQGS 63
 |||||
 Qy 11 SAIQGS 16

RESULT 47
 ID VGH_BPPHK STANDARD; PRT; 332 AA.
 AC P03649;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE MINOR SPIKE PROTEIN (H PROTEIN) (PILOT PROTEIN).
 GN H.
 OS Bacteriophage phi-K.
 OC Viruses; ssDNA viruses; Microviridae; Microvirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kodaira K.-I., Tadokoro H., Suzuki K., Oki M., Manto S., Taketo A.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-42 FROM N.A.
 RX MEDLINE; 80049950.
 RA Sims J., Capon D., Dressler D.;
 RT "dnaG (primase)-dependent origins of DNA replication. Nucleotide
 RT sequences of the negative strand initiation sites of bacteriophages
 RT St-1, phi K, and alpha 3".
 RT J. Biol. Chem. 254:12613-12628(1979).
 CC -!- FUNCTION: MINOR SPIKE COMPONENT OF THE VIRAL SHELL. H PROTEIN IS
 CC INVOLVED IN THE EJECTION OF THE PHAGE DNA IN THE HOST AND IS
 CC INJECTED WITH THE DNA IN THE PERIPLASMIC SPACE OF THE HOST.
 CC -!- SUBUNIT: THE VIRION IS COMPOSED OF 60 COPIES EACH OF THE F, G, AND
 CC J PROTEINS, AND 12 COPIES OF THE H PROTEIN. THERE ARE 12 SPIKES
 CC WHICH ARE EACH COMPOSED OF 5 G AND ONE H PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X60323; CAA42893.1; -
 DR EMBL; M10726; AAA32365.1; -
 DR PIR; B04256; B04256.
 KW Coat protein.
 SQ SEQUENCE 332 AA; 35110 MW; C895959D1F9C5A7A CRC64;

Query Match 5.9%; Score 6; DB 1; Length 332;
 Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 60 SAIQGS 65
 |||||
 Qy 11 SAIQGS 16

RESULT 48
 ID VGH_BPG4 STANDARD; PRT; 337 AA.
 AC P03647;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE MINOR SPIKE PROTEIN (H PROTEIN) (PILOT PROTEIN).
 GN H.
 OS Bacteriophage G4.
 OC Viruses; ssDNA viruses; Microviridae; Microvirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 79053264.
 RA Godson G.N., Barrell B.G., Staden R., Fiddes J.C.;
 RT "Nucleotide sequence of bacteriophage G4 DNA.";
 RL Nature 276:236-247(1978).
 CC -!- FUNCTION: MINOR SPIKE COMPONENT OF THE VIRAL SHELL. H PROTEIN IS
 CC INVOLVED IN THE EJECTION OF THE PHAGE DNA IN THE HOST AND IS
 CC INJECTED WITH THE DNA IN THE PERIPLASMIC SPACE OF THE HOST.
 CC -!- SUBUNIT: THE VIRION IS COMPOSED OF 60 COPIES EACH OF THE F, G, AND
 CC J PROTEINS, AND 12 COPIES OF THE H PROTEIN. THERE ARE 12 SPIKES
 CC WHICH ARE EACH COMPOSED OF 5 G AND ONE H PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; J02454; AAA32325.1; -
 DR EMBL; V00657; CAA24021.1; -
 DR PIR; A04255; ZHBPG4.
 KW Coat protein.
 SQ SEQUENCE 337 AA; 35706 MW; 3D8E94F2E21CB9AB CRC64;

Query Match 5.9%; Score 6; DB 1; Length 337;
 Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 54 SAIQGS 59
 |||||
 Qy 11 SAIQGS 16

RESULT 49
 ID YG5F_YEAST STANDARD; PRT; 366 AA.
 AC P53320;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE PUTATIVE MITOCHONDRIAL CARRIER YGR257C.
 GN YGR257C OR G9175.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
 OC Saccharomycetaceae; Saccharomycetes.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN-S288C;
 RX MEDLINE; 97279233.
 RA Mazzoni C., Ruzzi M., Rinaldi T., Solinas F., Montebove F.,
 RA Frontali L.;
 RT "Sequence analysis of a 10.5 kb DNA fragment from the yeast
 RT chromosome VII reveals the presence of three new open reading frames
 RT and of a tRNA^{Thr} gene.";
 RL Yeast 13:369-372(1997).
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL

CC INNER MEMBRANE (POTENTIAL).
CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X99228; CAA67613.1; -;
DR EMBL; JF3042; CAA97286.1; -;
DR PFAM; PF00153; mito_carr; 2;
DR PROSITE; PS00215; MITOCH_CARRIER; 1;
KW Hypothetical protein: Mitochondrion; Inner membrane; Repeat;
KW Transmembrane; Transport.
FT TRANSMEM 17 36 POTENTIAL.
FT TRANSMEM 124 140 POTENTIAL.
FT TRANSMEM 162 182 POTENTIAL.
FT TRANSMEM 268 286 POTENTIAL.
SQ SEQUENCE 366 AA; 40763 MW; B0358B6E818CB1E CRC64;

Query Match 5.9%; Score 6; DB 1; Length 366;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 GSVLTS 27
QY 15 GSVLTS 20
|||||

RESULT 50
ID AAAD_HUMAN STANDARD; PRT; 398 AA.
AC P22760;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE ARYLACETAMIDE DEACETYLASE (EC 3.1.1.-) (AADAC).
GN DAC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
[1]
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=LIVER;
RX MEDLINE; 94342357.
RA Probst M.R., Beer M., Beer D., Jenoe P., Meyer U.A., Gasser R.;
RT "Human liver arylacetamide deacetylase. Molecular cloning of a novel
RT esterase involved in the metabolic activation of arylamine
RT carcinogens with high sequence similarity to hormone-sensitive
RT lipase.";
RL J Biol. Chem. 269:21650-21656(1994).
[2]
RN PARTIAL SEQUENCE, AND CHARACTERIZATION.
RP TISSUE=LIVER;
RX MEDLINE; 91254316.
RA Probst M.R., Jenoe P., Meyer U.A.;
RT Purification and characterization of a human liver arylacetamide
RT deacetylase.";
RL Biochem. Biophys. Res. Commun. 177:453-459(1991).
[1]
CC -!- FUNCTION: ARYLACETAMIDE DEACETYLATION IS AN IMPORTANT ENZYME
CC ACTIVITY IN THE METABOLIC ACTIVATION OF ARYLAMINE SUBSTRATES TO
CC ULTIMATE CARCINOGENS.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MICROSOMAL
CC (POTENTIAL).
CC -!- TISSUE SPECIFICITY: LIVER.
CC -!- SIMILARITY: BELONGS TO THE "GDXG" FAMILY OF LIPOLYTIC ENZYMES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L32179; AAA35551.1; -;
DR MIN; 600338; -;
DR PFAM; PF00135; Coesterase; 1;
DR PROSITE; PS01174; LIPASE_GDXG_SER; 1;
KW Hydrolase; Transmembrane; Microsome; Signal-anchor.
FT INIT_MET 0 0
FT TRANSMEM 4 24 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT ACT_SITE 110 110 (POTENTIAL).
FT ACT_SITE 188 188 POTENTIAL.
FT CONFLICT 2 2 R -> M (IN REF. 2).
SQ SEQUENCE 398 AA; 45540 MW; 201FD6AC1700AD0C CRC64;

Query Match 5.9%; Score 6; DB 1; Length 398;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 302 GSSELA 307
QY 65 GSSELA 70
|||||

RESULT 51
ID PAP3_XENLA STANDARD; PRT; 400 AA.
AC P51006;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE POLY(A) POLYMERASE TYPE 3 (EC 2.7.7.19) (PAP) (POLYNUCLEOTIDE
DE ADENYLTRANSFERASE) (FRAGMENT).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=OVARY;
RX MEDLINE; 96079940.
RA Ballantyne S., Bilger A., Astrom J., Virtanen A., Wickens M.;
RT "Poly (A) polymerases in the nucleus and cytoplasm of frog oocytes:
RT dynamic changes during oocyte maturation and early development.";
RL RNA 1:64-78(1995).
CC -!- FUNCTION: POLYMERASE THAT CREATES THE 3' POLY(A) TAIL OF MRNA'S.
CC MAY ACQUIRE SPECIFICITY THROUGH INTERACTION WITH A CLEAVAGE AND
CC POLYADENYLATION FACTOR (CPSF).
CC -!- CATALYTIC ACTIVITY: N ATP + (NUCLEOTIDE)(M) - N PYROPHOSPHATE +
CC (NUCLEOTIDE)(M+N).
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U19975; AAC59747.1; -;
DR PFAM; PF01909; NTP_transf_2; 1.
KW mRNA processing; Transferase; Transcription; RNA-binding;
KW Nuclear protein.
FT ACT_SITE 110 110 BY SIMILARITY.
FT ACT_SITE 112 112 BY SIMILARITY.
FT ACT_SITE 164 164 BY SIMILARITY.
FT DOMAIN 382 390 NUCLEAR LOCALIZATION SIGNAL 1 (BY
FT NON_TER 400 400 SIMILARITY).

RA Glaser P., Danchin A.;
 RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE.
 CC -1- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + GLYCINE +
 CC H(2)O = TETRAHYDROFOLATE + L-SERINE.
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC -1- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,
 CC HORMONES AND OTHER COMPONENTS.
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SIMILARITY: BELONGS TO THE SHMT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; 238002; CAA86110.1; -;
 CC EMBL; 299122; CAB15707.1; -;
 CC SUBTILIST; BG10944; GLYA.
 CC PFAM; PF00464; SHMT; 1.
 CC PROSITE; PS00096; SHMT; 1.
 CC TRANSFERASE; PYRIDOXAL PHOSPHATE: One-carbon metabolism.
 CC FT BINDING 226 226 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC SEQUENCE 415 AA; 45489 MW; A3AD2F7C40AB14A1 CRC64;
 CC
 CC Query Match 5.9%; Score 6; DB 1; Length 415;
 CC Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC Db 43 QGSVLT 48
 CC | | | | |
 CC Qy 14 QGSVLT 19
 CC
 CC RESULT 55
 CC ID GLYA_STRCO STANDARD; PRT; 420 AA.
 CC AC O86565;
 CC DT 15-FEB-2000 (Rel. 39, Created)
 CC DT 15-FEB-2000 (Rel. 39, Last sequence update)
 CC DT 15-FEB-2000 (Rel. 39, Last annotation update)
 CC DE SERINE HYDROXYMETHYLTRANSFERASE (EC 2.1.2.1) (SERINE METHYLASE)
 CC (SHMT).
 CC GN GLYA OR SC2A11.04C.
 CC OS Streptomyces coelicolor.
 CC OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN=A3(2);
 CC RA Murphy L., Harris D., Parkhill J., Barrell B.G., Rajandream M.A.;
 CC Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE.
 CC -1- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + GLYCINE +
 CC H(2)O = TETRAHYDROFOLATE + L-SERINE.
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC -1- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,
 CC HORMONES AND OTHER COMPONENTS.
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SHMT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AL031184; CAA20173.1; -;
 DR PFAM; PF00464; SHMT; 1.
 DR PROSITE; PS00096; SHMT; 1.
 KW Transferase; Pyridoxal phosphate: One-carbon metabolism.
 FT BINDING 229 229 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 SQ SEQUENCE 420 AA; 44776 MW; 6BFD07E89AA8BFC CRC64;
 CC
 CC Query Match 5.9%; Score 6; DB 1; Length 420;
 CC Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC Db 47 QGSVLT 52
 CC | | | | |
 CC Qy 14 QGSVLT 19
 CC
 CC RESULT 56
 CC ID GLA2_MYCTU STANDARD; PRT; 425 AA.
 CC AC O53615;
 CC DT 15-FEB-2000 (Rel. 39, Created)
 CC DT 15-FEB-2000 (Rel. 39, Last sequence update)
 CC DT 15-FEB-2000 (Rel. 39, Last annotation update)
 CC DE SERINE HYDROXYMETHYLTRANSFERASE 2 (EC 2.1.2.1) (SERINE METHYLASE 2)
 CC (SHMT 2).
 CC GN GLYA2 OR RV0070 OR MTV030.13C.
 CC OS Mycobacterium tuberculosis.
 CC OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN=H37RV;
 CC MEDLINE; 98295987.
 CC RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 CC Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 CC Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 CC Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 CC Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 CC Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 CC Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
 CC Taylor K., Whitehead S., Barrell B.G.;
 CC "Deciphering the biology of Mycobacterium tuberculosis from the
 CC complete genome sequence."
 CC Nature 393:537-544(1998).
 CC -1- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE.
 CC -1- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + GLYCINE +
 CC H(2)O = TETRAHYDROFOLATE + L-SERINE.
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC -1- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,
 CC HORMONES AND OTHER COMPONENTS.
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SHMT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AL021428; CAA16251.1; -;
 CC TUBERCULIST; RV0070; -;
 CC PFAM; PF00464; SHMT; 1.
 CC PROSITE; PS00096; SHMT; 1.
 KW Transferase; Pyridoxal phosphate: One-carbon metabolism.
 FT BINDING 230 230 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 SQ SEQUENCE 425 AA; 45525 MW; 224D195C1D8BF680 CRC64;
 CC
 CC Query Match 5.9%; Score 6; DB 1; Length 425;
 CC Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC Db 47 QGSVLT 52


```

QY      14 QGSVLT 19
      |||||
RESULT  57
ID      GLAL_MYCTU      STANDARD;      PRT;      426 AA.
AC      053441.
DT      15-FEB-2000 (Rel. 39, Created)
DT      15-FEB-2000 (Rel. 39, Last sequence update)
DT      15-FEB-2000 (Rel. 39, Last annotation update)
DE      SERINE HYDROXYMETHYLTRANSFERASE 1 (EC 2.1.2.1) (SERINE METHYLASE 1)
DE      (SHMT 1)
GN      GLYAL OR GLYA OR RV1093 OR MV017.46.
OS      Mycobacterium tuberculosis.
OC      Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC      Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Simmonds M.N., Badcock K., James K.D., Parkhill J., Barrell B.G.,
RA      Rajandream M.A.;
RL      Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC      -!- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE.
CC      -!- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + GLYCINE +
CC      H(2)O = TETRAHYDROFOLATE + L-SERINE.
CC      -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC      -!- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,
CC      HORMONES AND OTHER COMPONENTS.
CC      -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC      -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC      -!- SIMILARITY: BELONGS TO THE SHMT FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; AL049491; CAB39828.1;
DR      PROSITE; PS00096; SHMT; FALSE_NEG.
KW      Transferase; Pyridoxal phosphate; One-carbon metabolism.
FT      BINDING 227 227 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ      SEQUENCE 426 AA; 45224 MW; 27783E2328AF2C98 CRC64;

Query Match      5.9%; Score 6; DB 1; Length 426;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches      6; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

Db      44 QGSVLT 49
      |||||
QY      14 QGSVLT 19

RESULT  59
ID      GLYA_SYNY3      STANDARD;      PRT;      427 AA.
AC      P77962.
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      15-FEB-2000 (Rel. 39, Last annotation update)
DE      SERINE HYDROXYMETHYLTRANSFERASE (EC 2.1.2.1) (SERINE METHYLASE)
DE      (SHMT).
GN      GLYA OR SLL1931.
OS      Synecocystis sp. (strain PCC 6803).
OC      Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
RN      [1]
RP      SEQUENCE FROM N.A.
RA      MEDLINE; 97061201.
RA      Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA      Miyajima N., Hirose M., Sugita M., Sasamoto S., Kimura T.,
RA      Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA      Shimp S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA      Tabata S.;
RT      "Sequence analysis of the genome of the unicellular cyanobacterium
RT      Synecocystis sp. strain PCC6803. II. Sequence determination of the
RT      entire genome and assignment of potential protein-coding regions.";
RL      DNA Res. 3:109-136(1996).
CC      -!- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE.
CC      -!- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + GLYCINE +
CC      H(2)O = TETRAHYDROFOLATE + L-SERINE.
CC      -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC      -!- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,
CC      HORMONES AND OTHER COMPONENTS.
CC      -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).

```

CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE SHMT FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D90903; BAAL7124.1; -.
CC PFAM; PF00464; SHMT; 1.
DR PROSITE; PS00096; SHMT; 1.
KW Transferase; Pyridoxal phosphate; One-carbon metabolism.
FT BINDING 231 231 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 427 AA; 46259 MW; B88AE834E2FA6045 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 427;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

48 QGSVLT 53
|||||
Qy 14 QGSVLT 19

RESULT 60
ID CD44_MESAU STANDARD; PRT; 431 AA.
AC Q60522; Q60523;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE CD44 ANTIGEN PRECURSOR (PHAGOCYTIC GLYCOPROTEIN I) (PGP-1) (HUTCH-I)
DE (EXTRACELLULAR MATRIX RECEPTOR-III) (ECMR-III) (GP90 LYMPHOCYTE
DE HOMING/ADHESION RECEPTOR) (HERMES ANTIGEN) (HYALURONATE RECEPTOR)
DE (HEPARAN SULFATE PROTEOGLYCAN) (HAMI ANTIGEN).
GN CD44.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-LVG; TISSUE-ALVEOLAR MACROPHAGE;
RA Paulauskis J.D., Kobzik L., Gerard C., Katler M., Godleski J.J.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases
CC -!- FUNCTION: MAIN CELL SURFACE RECEPTOR FOR HYALURONATE. ADHESION TO
CC MUCOSAL HIGH ENDOTHELIAL VENULE AND TO TYPES I AND VI COLLAGEN.
CC PROBABLY INVOLVED IN MATRIX ADHESION, LYMPHOCYTE ACTIVATION AND
CC LYMPH NODE HOMING.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- ALTERNATIVE PRODUCTS: THERE EXIST AT LEAST 2 ISOFORMS DUE TO
CC ALTERNATIVE SPLICING OF THE SAME GENE.
CC -!- PTM: EXTENSIVELY MODIFIED INCLUDING N- AND O-LINKED GLYCOSYLATION,
CC ADDITION OF THE GLYCOSAMINOGLYCAN CHONDROITIN SULFATE, OF SULFATE,
CC OF PHOSPHATE TO CYTOPLASMIC DOMAIN SERINE RESIDUES (BY
CC SIMILARITY).
CC -!- SIMILARITY: CONTAINS 1 LINK DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U10880; AAC13767.1; -.
CC EMBL; U10881; AAA19316.1; -.
DR HSSP; P98066; ITSG.
DR PFAM; PF00193; Xlink; 1.
DR PRINTS; PR00658; CD44.

DR PROSITE; PS01241; LINK; 1.
KW Cell adhesion; Transmembrane; Glycoprotein; Phosphorylation;
KW Receptor; Proteoglycan; Sulfatation; Signal; Alternative splicing.
FT SIGNAL 1 20
FT CHAIN 21 431
FT DOMAIN 21 339
FT TRANSMEM 340 360
FT DOMAIN 361 431
FT DOMAIN 50 121
FT DISULFID 55 120
FT DISULFID 79 99
FT CARBOHYD 27 27
FT CARBOHYD 59 59
FT CARBOHYD 102 102
FT CARBOHYD 112 112
FT CARBOHYD 122 122
FT CARBOHYD 174 174
FT CARBOHYD 256 256
FT CARBOHYD 325 325
FT VARSPLIC 222 291
TRSGKGQRRGGGLPKDATTSLGVTTHYPETMENGTLTPV
TPAKTGVGEETVVAEDSNFVDSLPG -> R (IN AN
ISOFORM).
SQ SEQUENCE 431 AA; 46807 MW; 4300262E0C6BEA6A CRC64;

Query Match 5.9%; Score 6; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

284 NVDGSL 289
|||||
Qy 42 NVDGSL 47

RESULT 61
ID MOTC_RHIME STANDARD; PRT; 434 AA.
AC Q52963;
DT 15-FEB-2000 (Rel. 39, Created)
DT 15-FEB-2000 (Rel. 39, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE CHEMOTAXIS MOTC PROTEIN PRECURSOR (MOTILITY PROTEIN C).
GN MOTC.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RU11/001;
RX MEDLINE; 97474284.
RA Platzner J., Sterr W., Hausmann M., Schmitt R.;
RT "Three genes of a motility operon and their role in flagellar rotary
RT speed variation in Rhizobium meliloti.";
RL J. Bacteriol. 179:6391-6399(1997).
CC -!- FUNCTION: REQUIRED FOR THE ROTATION OF THE FLAGELLAR MOTOR. MIGHT
CC CONTROL THE ENERGY FLUX OR COUPLING THAT DRIVES FLAGELLAR
CC ROTATION.
CC -!- SUBCELLULAR LOCATION: PERIPLASMIC.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L49337; AAB81409.1; -.
DR Chemotaxis; Flagella; Flagellar rotation; Periplasmic; Signal.
FT SIGNAL 1 15
FT CHAIN 16 434
FT SEQUENCE 434 AA; 46542 MW; EB89B388007A0B71 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 434;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 369 SELAAE 374
|||||

QY 67 SELAAE 72

RESULT 62
ID GLYA_METEX STANDARD; PRT; 434 AA.
AC P50435;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE SERINE HYDROXYMETHYLTRANSFERASE (EC 2.1.2.1) (SERINE METHYLASE)
DE (SHMT).
GN GLYA.

OS Methylobacterium extorquens.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Methylobacterium group; Methylobacterium.

[1]
SEQUENCE FROM N.A.
RC STRAIN-AM1 / NCIB 9133;
RX MEDLINE: 95050239.
RA Chistoserdova L.V., Lidstrom M.E.;
RT "Genetics of the serine cycle in Methylobacterium extorquens AM1:
RT cloning, sequence, mutation, and physiological effect of glyA, the
RT gene for serine hydroxymethyltransferase.";
RL J. Bacteriol. 176:6759-6762(1994).
CC -!- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE. MAY BE REQUIRED
CC FOR BOTH C1 AND C2 METABOLISM.
CC -!- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + GLYCINE +
CC H(2)O = TETRAHYDROFOLATE + L-SERINE.
CC -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC -!- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,
CC HORMONES AND OTHER COMPONENTS.
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE SHMT FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL: L33463; AAA64456.1; -;
PFAM: PF00464; SHMT; 1.
DR PROSITE; PS00096; SHMT; 1.
KW Transferase; Pyridoxal phosphate; One-carbon metabolism.
FT BINDING 242 242 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 434 AA; 46305 MW; EC9599B1E1AE44BC CRC64;

Query Match 5.9%; Score 6; DB 1; Length 434;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 59 QGSVLT 64
|||||

QY 14 QGSVLT 19

RESULT 63
ID SLAP_LACAC STANDARD; PRT; 444 AA.
AC P35829;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE S-LAYER PROTEIN PRECURSOR (SURFACE LAYER PROTEIN) (SA-PROTEIN).
GN SLPA.

OS Lactobacillus acidophilus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;

OC Lactobacillus.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 343-351 AND 440-444.
RC STRAIN-ATCC 4356;
RX MEDLINE: 94012467.
RA Boot H.J., Kolen C.P.A.M., van Noort J.M., Pouwels P.H.;
RT "S-layer protein of Lactobacillus acidophilus ATCC 4356:
RT purification, expression in Escherichia coli, and nucleotide sequence
RT of the corresponding gene.";
RL J. Bacteriol. 175:6089-6096(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 4356;
RX MEDLINE: 96099308.
RA Boot H.J., Kolen C.P.A.M., Pouwels P.H.;
RT "Identification, cloning, and nucleotide sequence of a silent S-layer
RT protein gene of Lactobacillus acidophilus ATCC 4356 which has
RT extensive similarity with the S-layer protein gene of this species.";
RL J. Bacteriol. 177:7222-7230(1995).
CC -!- FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY
CC OF PROTEINS WHICH COAT THE SURFACE OF BACTERIA.
CC -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIA IS COVERED BY A
CC S-LAYER WITH HEXAGONAL SYMMETRY.
CC -!- SIMILARITY: SOME, TO THE S-LAYER PROTEIN OF L.BREVIS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL: X71412; CAA50535.1; -;
DR EMBL: X89375; CAA61560.1; -;
DR PIR: A36924; A36924.
KW Signal; Glycoprotein; Cell wall; S-layer.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 444 S-LAYER PROTEIN.
SQ SEQUENCE 444 AA; 46570 MW; 2050732F89099161 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 444;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 377 ANIDGT 382
|||||

QY 92 ANIDGT 97

RESULT 64
ID VANS_ENTFA STANDARD; PRT; 447 AA.
AC Q47745;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE SENSOR PROTEIN VANSB (EC 2.7.3.-) (VANCOMYCIN B-TYPE RESISTANCE
DE PROTEIN VANSB) (VANCOMYCIN HISTIDINE PROTEIN KINASE).
GN VANSB.

OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;
OC Enterococcus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-V583;
RX MEDLINE: 96200099.
RA Evers S., Courvalin P.;
RT "Regulation of VanB-type vancomycin resistance gene expression by the
RT VanS(B)-VanR (B) two-component regulatory system in Enterococcus
RT faecalis V583.";
RL J. Bacteriol. 178:1302-1309(1996).
CC -!- FUNCTION: MEMBER OF THE TWO-COMPONENT REGULATORY SYSTEM VANS/VANR.
CC ACTIVATES THE TRANSCRIPTION OF VANH, VANA AND VANX IN RESPONSE TO

```

CC CC VANCOMYCIN WHICH RESULTS IN VANCOMYCIN RESISTANCE. VANS ACTIVATES
CC CC VANR BY PHOSPHORYLATION.
CC CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC CC -1- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE
CC CC KINASES.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; U35369; AAB05623.1; -.
CC CC PFAM; PF00672; DUF5.1;
CC CC PFAM; PF00512; signal.1;
CC CC Sensory transduction; Transferase; Kinase; Transmembrane; Cell wall;
CC CC Antibiotic resistance; Phosphorylation.
CC CC TRANSMEM 10 30 POTENTIAL.
CC CC TRANSMEM 137 155
CC CC MOD_RES 233 233
CC CC SEQUENCE 447 AA; 50111 MW; 6471F8A63C7498A2 CRC64;
CC CC -----
CC CC Query Match 5.9%; Score 6; DB 1; Length 447;
CC CC Best Local Similarity 100.0%; Pred. No. 8.75e+00;
CC CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC CC -----
Db 30 AQQFVS 35
CC | | | | |
QY 77 AQQFVS 82
CC -----
RESULT 65
ID AMYB_BACFI STANDARD; PRT; 468 AA.
AC P96513;
AT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE BETA-AMYLASE PRECURSOR (EC 3.2.1.2) (1,4-ALPHA-D-GLUCAN
DE MALTOHYDROLASE) (FRAGMENT).
DE Bacillus firmus.
OS Bacillus firmus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
RN [1]
RP SEQUENCE FROM N.A.
RA Chen W., He W.B.;
RT "Nucleotide sequence and characteristics of beta-amylase gene from
RT Bacillus firmus.";
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CATALYZES THE LIBERATION OF MALTOSE FROM 1,4-ALPHA-D-
CC GLUCANS.
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC LINKAGES IN
CC POLYSACCHARIDES SO AS TO REMOVE SUCCESSIVE MALTOSE UNITS FROM THE
CC NON-REDUCING ENDS OF THE CHAINS.
CC -1- SIMILARITY: BELONGS TO FAMILY 14 OF GLYCOSYL HYDROLASES
CC (BETA-AMYLASES).
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; AB000264; BAA19075.1; -.
CC CC PFAM; PF01373; Glyco_hydro.14; 1.
CC CC PROSITE; PS00506; BETA_AMYLASE_1; 1.
CC CC PROSITE; PS00679; BETA_AMYLASE_2; 1.
CC CC hydrolase; Glycosidase; Polysaccharide degradation; Signal.
CC CC SIGNAL 1 36 POTENTIAL.
CC CC CHAIN 37 >468 BETA-AMYLASE.
CC CC -----

```

```

RESULT 67
ID COX1_PLA8E STANDARD; PRT; 476 AA.
AC O99252;
DT 15-FEB-2000 (Rel. 39, Created)
DT 15-FEB-2000 (Rel. 39, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE I (EC 1.9.3.1).
GN COI.
OS Plasmodium berghei.
OG Mitochondrion.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
RN [1]
RP SEQUENCE FROM N.A.
RA Tan T.M.C., Noviyanti R., Syafrudi N., Marzuki S., Ting R.C.Y.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY
CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNITS 1-
3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. CO I IS THE
CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN
CYTOCHROME C ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2
AND HEME A OF SUBUNIT 1 TO THE BIMETALLIC CENTER FORMED BY HEME A3
AND COPPER B.
CC -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
4 FERRICYTOCHROME C.
CC -!- PATHWAY: TERMINAL STEP IN THE RESPIRATORY CHAIN.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN, MITOCHONDRIAL
INNER MEMBRANE. CONTAINS 12 POTENTIAL TRANSMEMBRANE DOMAINS.
CC -!- SIMILARIY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; AF014115; AAD01525.1; ALT_SEQ.
DR PFAM; PF00115; COX1; 1.
DR PRINTS; PR01165; CYCOXIDASE1.
DR PROSITE; PS00077; COX1; 1.
KW Oxidoreductase; Heme; Copper; Mitochondrion; Transmembrane;
Respiratory chain; Inner membrane.
FT METAL 66 66 IRON (HEME A) (PROBABLE).
FT METAL 246 246 COPPER B (PROBABLE).
FT METAL 250 250 COPPER B (PROBABLE).
FT METAL 295 295 COPPER B (PROBABLE).
FT METAL 296 296 COPPER B (PROBABLE).
FT METAL 382 382 IRON (HEME A3) (PROBABLE).
FT METAL 384 384 IRON (HEME A) (PROBABLE).
SQ SEQUENCE 476 AA; 52728 MW; 690160EE61091C0 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 476;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 93 GSSELA 98
QY 65 GSSELA 70
|||||

RESULT 68
ID MKK1_YEAST STANDARD; PRT; 508 AA.
AC P32490;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE PROTEIN KINASE MKK1/SSP32 (EC 2.7.1.-).
GN MKK1 OR SSP32 OR YOR231W OR OS095.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;

Saccharomycetaceae; Saccharomyces.
[1]
RN RP SEQUENCE FROM N.A.
RA Irie K., Takase M., Lee K.S., Levin D.E., Araki H., Matsumoto K.,
Oshima Y.;
RT MKK1 and MKK2, which encode Saccharomyces cerevisiae
mitogen-activated protein kinase-kinase homologs, function in the
pathway mediated by protein kinase C.;
RL Mol. Cell. Biol. 13:3076-3083(1993).
[2]
RN RP SEQUENCE FROM N.A.
RC STRAIN-S288C / FY1679;
RA MEDLINE; 97127829.
RA Boyer J., Michaux G., Fairhead C., Gaillon L., Dujon B.;
RT "Sequence and analysis of a 26.9 kb fragment from chromosome XV of
the yeast Saccharomyces cerevisiae.";
RL Yeast 12:1575-1586(1996).
CC -!- FUNCTION: SERINE/THREONINE PROTEIN KINASE INVOLVED IN A SIGNAL
TRANSDUCTION PATHWAY THAT PLAY A ROLE IN YEAST CELL MORPHOGENESIS
AND CELL GROWTH. THIS PATHWAY SEEMS TO STARTS BY SMP3; THEN
INVOLVE THE KINASE PKC1 THAT MAY ACT ON THE BCK1 KINASE THAT THEN
PHOSPHORYLATES MKK1 AND MKK2 WHICH THEMSELVES PHOSPHORYLATE THE
MPK1 KINASE.
CC -!- FUNCTION: KINASE INVOLVED IN A SIGNAL TRANSDUCTION PATHWAY THAT
INVOLVE THE PROTEIN KINASES ENCODED BY PKC1, BCK1, AND MPK1.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
MAP KINASE KINASE SUBFAMILY
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; D13001; BAA02364.1; -.
DR EMBL; Z75139; CAA99451.1; -.
DR PIR; S30772; S30772.
DR PIR; A48069; A48069.
DR HSP; P11362; IFG1.
DR SGD; L0001117; MKK1.
DR PFAM; PF00069; pkinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
ATP-binding; Phosphorylation.
FT DOMAIN 221 488 PROTEIN KINASE.
FT NP_BIND 227 235 ATP (BY SIMILARITY).
FT BINDING 250 250 ATP (BY SIMILARITY).
FT ACT_SITE 349 349 BY SIMILARITY.
FT MOD_RES 377 377 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 381 381 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 508 AA; 56720 MW; 4FE62CDD185CDA2 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 508;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 196 IDLSNV 201
QY 34 IDLSNV 39
|||||

RESULT 69
ID CD39_HUMAN STANDARD; PRT; 510 AA.
AC P49961;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE VASCULAR ATP-DIPHOSPHOHYDROLASE (EC 3.6.1.5) (ATPDASE) (LYMPHOID CELL

```

DE ACTIVATION ANTIGEN) (CD39 ANTIGEN).
GN CD39.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95015846.
RA Maliszewski C.R., Delespesse G.J.T., Schoenborn M.A., Armitage R.J.,
RA Fanslow W.C., Nakajima T., Baker E., Sutherland G.R., Poindexter K.,
RA Birks C., Alpert A., Friend D., Gimpel S.D., Gayle R.B. III;
RT "The CD39 lymphoid cell activation antigen. Molecular cloning and
RT structural characterization."
RL J. Immunol. 153:3574-3583(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE-UMBILICAL VEIN;
RX MEDLINE; 97149443.
RA Robson S.C., Kaczmarek E., Siegel J.B., Candinas D., Kozlak K.,
RA Millan M., Hancock W.W., Bach F.H.;
RA "Loss of ATP diphosphohydrolase activity with endothelial cell
activation."
J. Exp. Med. 185:153-163(1997).
RN [3]
RP FUNCTION.
RX MEDLINE; 97115858.
RA Kaczmarek E., Kozlak K., Sevigny J., Siegel J.B., Anrather J.,
RA Beaudoin A.R., Bach F.H., Robson S.C.;
RT "Identification and characterization of CD39/vascular ATP
RT diphosphohydrolase."
RL J. Biol. Chem. 271:33116-33122(1996).
CC -1- CATALYTIC ACTIVITY: ATP + 2 H₂O = AMP + 2 PHOSPHATE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- TISSUE SPECIFICITY: EXPRESSED PRIMARILY ON ACTIVATED LYMPHOID
CC CELLS. ALSO EXPRESSED IN ENDOTHELIAL TISSUES.
CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.
CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD39 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd39.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S73813; AAB32152.1; -;
DR EMBL; U87967; AAB47572.1; -;
DR MIM; 601752; -;
PFAM; PF01150; GDAL_CD39; 1.
DR PROSITE; PS01236; GDAL_CD39_NTPASE; 1.
KW Hydrolase; Transmembrane; Antigen; Glycoprotein.
FT DOMAIN 1 16 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 17 38 POTENTIAL.
FT DOMAIN 39 478 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 479 499 POTENTIAL.
FT DOMAIN 500 510 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 73 73 POTENTIAL.
FT CARBOHYD 227 227 POTENTIAL.
FT CARBOHYD 292 292 POTENTIAL.
FT CARBOHYD 334 334 POTENTIAL.
FT CARBOHYD 371 371 POTENTIAL.
FT CARBOHYD 457 457 POTENTIAL.
SQ SEQUENCE 510 AA; 57964 MW; BAD87D2499649159 CRC64;
Query Match 5.9%; Score 6; DB 1; Length 510;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 186 LGKFSQ 191
Qy 1 LGKFSQ 6

DE ACTIVATION ANTIGEN) (CD39 ANTIGEN).
GN CD39.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95015846.
RA Maliszewski C.R., Delespesse G.J.T., Schoenborn M.A., Armitage R.J.,
RA Fanslow W.C., Nakajima T., Baker E., Sutherland G.R., Poindexter K.,
RA Birks C., Alpert A., Friend D., Gimpel S.D., Gayle R.B. III;
RT "The CD39 lymphoid cell activation antigen. Molecular cloning and
RT structural characterization."
RL J. Immunol. 153:3574-3583(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE-UMBILICAL VEIN;
RX MEDLINE; 97149443.
RA Robson S.C., Kaczmarek E., Siegel J.B., Candinas D., Kozlak K.,
RA Millan M., Hancock W.W., Bach F.H.;
RA "Loss of ATP diphosphohydrolase activity with endothelial cell
activation."
J. Exp. Med. 185:153-163(1997).
RN [3]
RP FUNCTION.
RX MEDLINE; 97115858.
RA Kaczmarek E., Kozlak K., Sevigny J., Siegel J.B., Anrather J.,
RA Beaudoin A.R., Bach F.H., Robson S.C.;
RT "Identification and characterization of CD39/vascular ATP
RT diphosphohydrolase."
RL J. Biol. Chem. 271:33116-33122(1996).
CC -1- CATALYTIC ACTIVITY: ATP + 2 H₂O = AMP + 2 PHOSPHATE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- TISSUE SPECIFICITY: EXPRESSED PRIMARILY ON ACTIVATED LYMPHOID
CC CELLS. ALSO EXPRESSED IN ENDOTHELIAL TISSUES.
CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.
CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD39 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd39.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S73813; AAB32152.1; -;
DR EMBL; U87967; AAB47572.1; -;
DR MIM; 601752; -;
PFAM; PF01150; GDAL_CD39; 1.
DR PROSITE; PS01236; GDAL_CD39_NTPASE; 1.
KW Hydrolase; Transmembrane; Antigen; Glycoprotein.
FT DOMAIN 1 16 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 17 38 POTENTIAL.
FT DOMAIN 39 478 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 479 499 POTENTIAL.
FT DOMAIN 500 510 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 73 73 POTENTIAL.
FT CARBOHYD 227 227 POTENTIAL.
FT CARBOHYD 292 292 POTENTIAL.
FT CARBOHYD 334 334 POTENTIAL.
FT CARBOHYD 371 371 POTENTIAL.
FT CARBOHYD 457 457 POTENTIAL.
SQ SEQUENCE 510 AA; 57964 MW; BAD87D2499649159 CRC64;
Query Match 5.9%; Score 6; DB 1; Length 510;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 186 LGKFSQ 191
Qy 1 LGKFSQ 6

RESULT 70
ID GDSL_YEAST STANDARD; PRT; 522 AA.
AC P41913;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE GDSL PROTEIN
GN GDSL OR YOR355W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC Saccharomycetaceae; Saccharomycetes.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-MB43-15C;
RX MEDLINE; 96353433.
RA Konopinska A., Szczesniak B., Boguta M.;
RT "Nucleotide sequence of the GDSL gene of Saccharomyces cerevisiae."
RL Yeast 11:1513-1518(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Delius H., Hebling U., Hofmann B.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INVOLVED IN NUCLEAR CONTROL OF MITOCHONDRIA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U18262; AAB17574.1; -;
DR EMBL; Z75263; CAA99684.1; -;
DR SGD; L0002580; GDSL.
KW Mitochondrion.
SQ SEQUENCE 522 AA; 57048 MW; 27F5B82703084715 CRC64;
Query Match 5.9%; Score 6; DB 1; Length 522;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 422 SSELAA 427
Qy 66 SSELAA 71

RESULT 71
ID UGT5_CAEEL STANDARD; PRT; 537 AA.
AC Q20086;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PUTATIVE UDP-GLUCURONOSYLTRANSFERASE UGT5 PRECURSOR (EC 2.4.1.17)
GN (UDPGT).
DE UGT5 OR F35H8.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-BRISTOL N2;
RX BERKS M.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: UDP-GLUCURONATE + ACCEPTOR = UDP + ACCEPTOR
CC BETA-D-GLUCURONOSIDE.
CC -1- SIMILARITY: BELONGS TO THE UDP-GLYCOSYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; Z36752; CAA85328.1; -
 WORKPEP; F35H8.6; C601575.
 PFAM; PF00201; UDPGT; 1.
 PROSITE; PS00375; UDPGT; FALSE_NEG.
 KW Hypothetical protein; Transferase; Glycosyltransferase; Glycoprotein;
 KW Transmembrane; Signal; Multigene family.
 FT SIGNAL 1 16
 CHAIN 17 537
 PUTATIVE UDP-GLUCURONOSYLTRANSFERASE
 UGT5.
 TRANSMEM 486 506
 CARBOHYD 88 88
 FT CARBOHYD 243 243
 POTENTIAL.
 FT CARBOHYD 336 336
 POTENTIAL.
 SEQUENCE 537 AA; 61641 MW; BF437944EBF5EED CRC64;

Query Match 5.9%; Score 6; DB 1; Length 537;
 Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 370 LNSVIE 375
 Qy 36 LNSVIE 41

RESULT 72

ID OPPIA_SALTY STANDARD; PRT; 542 AA.
 AC P06202;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE PERIPLASMIC OLIGOPEPTIDE-BINDING PROTEIN PRECURSOR.
 GN OPPIA.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Salmonella.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LT2;
 RX MEDLINE; 88011222.
 RA Hiles I.D., Gallagher M.P., Jamieson D.J., Higgins C.F.;
 RT "Molecular characterization of the oligopeptide permease of
 Salmonella typhimurium".
 J. Mol. Biol. 195:125-142(1987).
 [2]

RP SEQUENCE FROM N.A.
 RX MEDLINE; 86274740.
 RA Hiles I.D., Higgins C.F.;
 RT "Peptide uptake by Salmonella typhimurium. The periplasmic
 oligopeptide-binding protein".
 Eur. J. Biochem. 158:561-567(1986).
 [3]
 RN X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
 RP MEDLINE; 94261830.
 RX Tame J.R.H., Murshudov G.N., Dodson E.J., Neil T.K., Dodson G.G.,
 RA Higgins C.F., Wilkinson A.J.;
 RT "The structural basis of sequence-independent peptide binding by Oppa
 protein".
 Science 264:1578-1581(1994).
 [4]
 RN X-RAY CRYSTALLOGRAPHY (1.4 ANGSTROMS).
 RP MEDLINE; 96363676.
 RX Tame J.R.H., Murshudov G.N., Murshudov G.N., Higgins C.F.,
 RA Wilkinson A.J.;
 RT "The crystal structures of the oligopeptide-binding protein Oppa
 complexed with tripeptide and tetrapeptide ligands".
 Structure 3:1395-1406(1995).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).

RA Davies T.G., Tame J.R.H.;
 RL Submitted (NOV-1998) to the PDB data bank.
 CC FUNCTION: THIS PROTEIN IS A COMPONENT OF THE OLIGOPEPTIDE
 CC PERMEASE, A BINDING PROTEIN-DEPENDENT TRANSPORT SYSTEM, IT BINDS
 CC PEPTIDES UP TO FIVE AMINO ACIDS LONG WITH HIGH AFFINITY.
 CC SUBCELLULAR LOCATION: PERIPLASMIC.
 CC SIMILARITY: BELONGS TO THE BACTERIAL EXTRACELLULAR SOLUTE-BINDING
 CC PROTEIN FAMILY 5.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL; X04194; CAA27785.1; -
 EMBL; X05491; CAA29039.1; -
 PIR; A25011; QREBOA.
 DR PDB; 1OLA; 31-JUL-94.
 DR PDB; 2OLB; 29-JAN-96.
 DR PDB; 1OLC; 29-JAN-96.
 DR PDB; 1JET; 15-MAY-97.
 DR PDB; 1JEU; 15-MAY-97.
 DR PDB; 1JEV; 15-MAY-97.
 DR PDB; 1RKM; 29-JUL-97.
 DR PDB; 2RKM; 29-JUL-97.
 DR PDB; 1B05; 25-NOV-98.
 DR PDB; 1B0H; 18-NOV-98.
 DR PDB; 1B1H; 18-NOV-98.
 DR PDB; 1B2H; 18-NOV-98.
 DR PDB; 1B3F; 16-DEC-98.
 DR PDB; 1B3G; 16-DEC-98.
 DR PDB; 1B3H; 18-NOV-98.
 DR PDB; 1B3L; 16-DEC-98.
 DR PDB; 1B4H; 18-NOV-98.
 DR PDB; 1B42; 13-JAN-99.
 DR PDB; 1B5H; 18-NOV-98.
 DR PDB; 1B5I; 13-JAN-99.
 DR PDB; 1B5J; 13-JAN-99.
 DR PDB; 1B6H; 18-NOV-98.
 DR PDB; 1B7H; 25-NOV-98.
 DR PDB; 1B9J; 22-FEB-99.
 DR PDB; 1B32; 23-DEC-98.
 DR PDB; 1B40; 13-JAN-99.
 DR PDB; 1B46; 13-JAN-99.
 DR PDB; 1B51; 20-JAN-99.
 DR PDB; 1B52; 27-JAN-99.
 DR PDB; 1B58; 20-JAN-99.
 DR STYGENE; SG10267; OPPIA.
 DR PFAM; PF00496; SBP_bac.5; 1.
 DR PROSITE; PS01040; SBP_BACTERIAL_5; 1.
 KW Peptide transport; Transport; Periplasmic; Signal; 3D-structure.

FT SIGNAL 1 25
 CHAIN 26 542
 FT DISULFID 296 442
 FT SEQUENCE 542 AA; 61261 MW; 35F0C8E16B8348AA CRC64;
 Query Match 5.9%; Score 6; DB 1; Length 542;
 Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 142 HIANID 147
 Qy 90 HIANID 95

RESULT 73
 ID Y37B_MYCLE STANDARD; PRT; 558 AA.
 AC Q49755;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-FEB-2000 (Rel. 39, Last annotation update)
 DE HYPOTHETICAL 61.5 KDA PROTEIN U1937B.
 GN U1937B OR B1937_FL_4.
 OS Mycobacterium leprae.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Mycobacteriaceae; Mycobacterium.
 CC Actinomycetales; Corynebacterineae; Mycobacterium.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Smith D.R., Robison K.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: STRONG, TO SYNECHOCYSTIS PCC 6803 SLL0335, SOME TO
 CC M. TUBERCULOSIS RV2567.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U00016; AAA17160.1; -
 CC Hypothetical protein.
 FT DOMAIN 538 554 GLN-RICH
 SQ SEQUENCE 558 AA; 61522 MW; 4472E500100ABFA6 CRC64;
 Query Match 5.9%; Score 6; DB 1; Length 558;
 Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 29 GGYNTS 34
 QY [1][1][1][1][1]
 QY 27 GGYNTS 32
 RESULT 74
 ID HEMA_TAPIL STANDARD; PRT; 566 AA.
 AC P13102;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
 DE HEMAGGLUTININ HA2 CHAIN].
 GN HA.
 OS Influenza A virus (strain A/Pilot whale/Maine/328/84).
 OS Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
 OC Influenza virus A and B group.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chambers T.M., Yamnikova S., Kawaoka Y., Lvov D.K., Webster R.G.;
 RT "Antigenic and molecular characterization of subtype H13
 RT hemagglutinin of influenza virus.";
 RL Virology 172:180-188(1989).
 CC -1- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
 CC CELL RECEPTORS AND FOR INITIATING INFECTION.
 CC -1- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMER IS FORMED BY TWO CHAINS
 CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M26091; AAA43215.1; -
 CC PIR; B32664; HMIVT2.
 CC HSP; P03437; 2HMG.
 CC PFAM; PF00509; Hemagglutinin; 1.
 CC PRINTS; PR00329; HEMAGGLUTIN12.
 CC PRINTS; PR00330; HEMAGGLUTIN1.
 CC PRINTS; PR00331; HEMAGGLUTIN2.

DR PRINTS; PR00331; HEMAGGLUTIN2.
 KW Envelope protein; Hemagglutinin; Glycoprotein; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 342 HEMAGGLUTININ HA1 CHAIN.
 FT CHAIN 344 566 HEMAGGLUTININ HA2 CHAIN.
 FT CARBOHYD 29 29 POTENTIAL.
 FT CARBOHYD 54 54 POTENTIAL.
 FT CARBOHYD 182 182 POTENTIAL.
 FT CARBOHYD 183 183 POTENTIAL.
 FT CARBOHYD 305 305 POTENTIAL.
 FT CARBOHYD 488 488 POTENTIAL.
 FT CARBOHYD 497 497 POTENTIAL.
 SQ SEQUENCE 566 AA; 63049 MW; 74FF5A9860B8E59E CRC64;
 Query Match 5.9%; Score 6; DB 1; Length 566;
 Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 45 TSSIDL 50
 QY [1][1][1][1][1]
 QY 31 TSSIDL 36
 RESULT 75
 ID HEMA_IAGU2 STANDARD; PRT; 566 AA.
 AC P13103;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HA1 CHAIN;
 DE HEMAGGLUTININ HA2 CHAIN].
 GN HA.
 OS Influenza A virus (strain A/Gull/Maryland/704/77).
 OS Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
 OC Influenza virus A and B group.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE; 91220697.
 RA Nobusawa E., Aoyama T., Kato H., Suzuki Y., Tateno Y., Nakajima K.;
 RT "Comparison of complete amino acid sequences and receptor-binding
 RT properties among 13 serotypes of hemagglutinins of influenza A
 RT viruses.";
 RL Virology 182:475-485(1991).
 CC [2]
 RN RN
 RP SEQUENCE FROM N.A.
 RA MEDLINE; 89370299.
 RA Chambers T.M., Yamnikova S., Kawaoka Y., Lvov D.K., Webster R.G.;
 RT "Antigenic and molecular characterization of subtype H13
 RT hemagglutinin of influenza virus.";
 RL Virology 172:180-188(1989).
 CC -1- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
 CC CELL RECEPTORS AND FOR INITIATING INFECTION.
 CC -1- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMER IS FORMED BY TWO CHAINS
 CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D90308; BAA14338.1; -
 CC EMBL; M26090; AAA43214.1; -
 CC PIR; C39987; HMIVT1.
 CC PIR; A32664; HMIVT1.
 CC PIR; G33157; G33157.
 CC HSP; P03437; 2HMG.
 CC PFAM; PF00509; Hemagglutinin; 1.
 CC PRINTS; PR00329; HEMAGGLUTIN12.
 CC PRINTS; PR00330; HEMAGGLUTIN1.
 CC PRINTS; PR00331; HEMAGGLUTIN2.

KW Envelope protein; Hemagglutinin; Glycoprotein; Signal.
FT SIGNAL 1 18
FT CHAIN 19 342 HEMAGGLUTININ HA1 CHAIN.
FT CHAIN 344 566 HEMAGGLUTININ HA2 CHAIN.
FT CARBOHYD 29 29 POTENTIAL.
FT CARBOHYD 54 54 POTENTIAL.
FT CARBOHYD 182 182 POTENTIAL.
FT CARBOHYD 183 183 POTENTIAL.
FT CARBOHYD 305 305 POTENTIAL.
FT CARBOHYD 488 488 POTENTIAL.
FT CARBOHYD 497 497 POTENTIAL.
FT CONFLICT 80 80 V -> G (IN REF. 2).
FT CONFLICT 172 172 T -> N (IN REF. 2).
SQ SEQUENCE 566 AA; 63294 MW; 0965B35148F2215F CRC64;

Query Match 5.9%; Score 6; DB 1; Length 566;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

45 TSSIDL 50
|||||
QY 31 TSSIDL 36

Search completed: Wed Aug 16 09:55:49 2000
Job time : 45 secs.

